DOCUMENT RESUME

ED 041 996 24 UD 010 514

AUTHOR Baldwin, Alfred L.; And Others
TITLE Cognitive Content for Mother-Child Interactions. Final Report.
INSTITUTION Cornell Univ., Ithaca, N.Y. Center for Research on Education.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-6-1341
PUB DATE May 70
GRANT OEG-0-9-610326-4480
NOTE 281p.

EDRS PRICE EDRS Price MF-$1.25 HC-$4.15
DESCRIPTORS Behavior Patterns, Childhood Attitudes, Cognitive Development, Cognitive Measurement, *Cognitive Processes, Individual Differences, *Interaction Process Analysis, *Interpersonal Relationship, Lower Class, Middle Class, Mother Attitudes, Nonverbal Communication, *Parent Child Relationship, Research Methodology, *Socialization, Stereotypes, Upper Class

ABSTRACT

This study details three methods developed during the course of an investigation for describing adult-child interaction: (1) the "VINEX" category system for coding the actual language of the adult and the child; (2) a coding system for describing nonverbal behavior; and, (3) "Interaction Language," for the use of an observer in narrating the adult-child interaction. The test materials were small samples of mother-child pairs from New York City--one from West Harlem (with a middle and lower class subsample), and the other from Washington Square (white upper middle class). The empirical findings of the study were as follows: (1) the general pattern of mother-child interaction was a responsive one--a give and take interaction which was not balanced, but at the same time not markedly one-sided; (2) changes in interaction with age appeared to be partly due to the child's cognitive development, increase of explanations, and increased grammatical complexity for example; (3) the differences between the West Harlem and the Washington Square samples, as well as the difference between the lower and middle class West Harlem sample appeared to be small; and, (4) commonly held stereotypes of the family interaction of Harlem children were not supported by any of the data of the study. Appended are manuals on the methodology developed, extensive tabulations of test results, and interview formats used. (RJ)

Final Report

Project No. 6-1341
Grant No. OEG 0-9-610326-4480

COGNITIVE CONTENT OF MOTHER-CHILD INTERACTIONS

Alfred L. Baldwin
Clara P. Baldwin

with

Milton Seegmiller
Sheldon Frank, M.D.
Boyce Ford
Paul Ward

Cornell University
Ithaca, New York 14850

May 1970

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

PREFACE

This is the final report of a research project which began in 1966. It has been funded under a Project Literacy grant from the Office of Education.

Our major objective throughout the research has been to develop reliable measures for studying the ways in which children acquire cognitive functions in naturalistic situations, and to utilize these measures to study this socialization process in the interactions of a variety of children and significant socialization agents.

At the time when the study began, there were many statements in the literature regarding the influence of mother-child interactions on the cognitive development of particularly our poor children in urban ghettoes. Yet there was little empirical data to support those assumptions. Our data collection has focussed on the study of a number of the interactions of boys and their mothers from West Harlem. The sample includes both middle and lower class Negro boys who were born in four hospitals in that area. We also have a sample of mothers and boys from the area bordering Washington Square in New York. Most of these families are middle-class and academic.

Because the work has been carried on over several years, during which time our major research unit was moved from New York University to Cornell University, there are a number of people in both New York and Ithaca whose loyal efforts have been crucial to the successful completion of the project.

This research has produced a massive logistics problem, thousands of utterances and an equal number of observers' sentences to record, transcribe, code, and analyze. There were times when the task seemed impossible. It has also had many moments of excitement, satisfaction, and the authors will be always grateful to their staff.

Bonni Seegmiller was with us throughout the entire project, first as a graduate assistant and then as a research associate. She has helped with the development of the measures, supervised the large data collection in Harlem in 1968-69, and has done a great deal of coding and data analysis.

Dr. Sheldon Frank has served as our linguistic consultant throughout the project, and he has supervised the training of coders and the analysis of the children's and mother's language utilizing the measure of syntactic complexity developed by Frank and Osser (1970). He has also been invaluable in our discussions of children's language development.

Milton Seegmiller has been in charge of the data collection during the fall of 1969, and has been particularly involved in the complexity analysis of the children's and mother's language.

Shirley Cohen, a graduate student in Psychology at New York University, did all of the initial development of the VINEX coding system, working for nearly a year from transcripts of half-hour interactions of a group of 4 four year old children and their teacher in two standard play situations.

Since our return to Cornell, Boyce Ford has been our research associate, and has contributed innumerable hours to the computer analysis of some 130,000 utterances. He has persistently fought and won a number of battles with the computer, and this report could never have been completed without his dedicated supervision of this part of the project.

Paul Ward, a fellow at the Center for Research in Education at Cornell, has contributed his creative knowledge of computer programming to the project, and it is primarily due to his assistance that we are well on the way to developing a program for analyzing the interactional language in such a way that we feel it will constitute an important break-through in the analysis of narrative records of naturalistic situations.

This kind of research has many details which must be attended to with care and accuracy. Accurate transcripts are the basis of the language analysis, and it is no easy job to find people who will do this work well. We have been fortunate in having a series of transcribers and typists who were not only skilled at their tasks, but dedicated enough to the research and the data to be interested in it as well as efficient. Kathy Woznicki, Elayne Barun, Mary Odum, Claire Browne, Shirley Hancy were five of these.

We've also had a series of research assistants and associates both at New York University and Cornell who have conscientiously and patiently carried out parts of data collection, the coding and analysis. Often it has been hard for them to see the total picture for the forest of detail of their particular task, but they have all effectively carried out the jobs that needed to be done, and in addition have contributed much to the research in the way of new ideas and clarification of old ones. Some of the assistants have been full-time, some graduare assistants and some Antioch co-op students. Working with us in New York were: Jared Keil, Ann Singer, Susan Feldman, Jan Drucker, George Green, Susan Blumenthal; at Cornell - Jane Hammacher, Bob Delestrada, Mari Peterson, Sharon Horner, Jean Simmons, Barbara Nelson, Jean Grossman have been our research associate and assistants.

Last, and most importantly, we would like to thank the parents and children who have helped us to learn more about mother-child interactions. We had approached research in New York with some trepidation, but have found the mothers and children a most delightful group of people to know. They have been interested in the research, reliable in meeting appointments, and cooperative with us. For some this has meant long trips from Harlem to New York University, and for 11 mothers it has meant staying with us for three years of research. Sincere thanks of the authors go to these people who have literally made the research possible.

There is little doubt that much of the cooperation which we received from some of the Harlem mothers was due to their association with Dr. Frank Palmer and his research group at the Harlem Training Center. This group has provided us with space to conduct the research and invaluable assistance in obtaining subjects and providing us with data from their own research.

Finally, our thanks go to our highly skillful and patient secretary, Mrs. Violet Shepardson, who has typed, re-typed, edited and assembled this report.

Alfred L. Baldwin
Clara P. Baldwin

Table of Contents

List of Tables in Body of Report

Chapter 1

INTRODUCTION AND BACKGROUND OF THE STUDY

It is a well established fact that children in some segments of our society, particularly in our inner cities, fare poorly in our educational system as it is presently conducted. For example, a report of reading achievement test scores for New York City as a whole and for each individual school was presented by the NEW YORK TIMES on February 15, 1970. This report was based on data prepared by Dr. Samuel McClellan, acting director of the Bureau of Educational Research of the New York City Board of Education. In 1969, for the city as a whole, second graders are only slightly behind the national norms, 55.3% as contrasted to the 50% U.S. norm line. By fourth grade the percentage falling below this norm has risen to 67.5 %. These percentages vary markedly from school to school, with the greatest discrepancies in general being in the poorer areas of the city. Of children entering high school a substantial group are functionally illiterate. They have sat through nine years of education without gaining any appreciable academic knowledge, and by the time they leave school, when they reach the earliest legal age, they have progressed no further.

Many different reasons have been suggested to account for this educational failure. Some people, particularly members of black community groups, attribute this failure to the school system itself. They insist that the predominately middle-class teacher assumes that poor black children are unteachable and never gives them a reasonable chance to learn. The fact that some "ghetto" schools have very good achievement test scores while other schools a few blocks away have poor achievement records suggests the partial validity of this contention. At the other extreme Jensen and others have suggested the possibility that there are genetic differences in intelligence between the races, and thus tacitly attribute the failure either to genetic deficit, or at least to the failure of the educational system to base its program upon the genetic endowments of children of other races than the dominant white race that forms the bulk of the school population. The history of earlier immigrant groups to the United States, who must have been as genetically different as the so called "Negro" population, hardly bears out any hypothesis of a genuine genetic deficit. After all very few blacks are even close to being genetically pure descendants of the imported African slaves of two hundred years ago. Most of them are the products of many inter-race unions. (Reed, 1969)

At the time we began this study in 1964, the most commonly accepted hypothesis was that the family life of the black children in Harlem was the cause of their educational difficulties. Broken homes, father absence, and poverty were hypothesized by the Moynihan report(1967) to result in an educational deficit that left the child unable to profit from the usual educational program. This hypothesis has led to massive expenditures of public money for compensatory education to provide the warmth, the interpersonal interaction, and the verbal stimulation of which the children in ghetto families are presumably deprived.

The history of this belief in the deprivation of the lower class children is interesting to follow. To hark back to an earlier era , many of us can remember when the lower class home was viewed as a free, warm, independence-giving environment, in contrast to the "up tight" middle class home with its rigid demands, its anxiety, and its high pressure for inhibition of natural impulses. (Davis & Havinghurst 1946) Gradually, however, these same homes -- even with some of the same presumed characteristics -- were viewed as depriving rather than free from pressure, as stultifying rather than free from anxiety.

As early as 1951 Milner reported that first grade children in a large southern city who were "low scorers" on a battery of language measures -- who turned out to be children from low SES homes as measured by Warner's index -- were subject to a variety of deprivations. In such families there was little or no conversation at meal times, partly because there were no regular meal times. A significant number report no conversation between the child and any other person between breakfast and going to school. Mothers in such homes also displayed less overt affection than mothers of high scoring children. High scoring children also had more books available to them and were read to by adults more frequently. All of these indices were derived from interviews with the mothers and children. This same article recommends as a remedial measure more verbal interaction between these children and significant adults (Milner, 1951).

There are many other references to the fact that lack of verbal interaction (particularly complex verbal interaction) in the home accounts for the need for intervention by outside agencies. Bernstein (1967) has probably stimulated much of this thinking through his findings that the working class mothers in England tend to use a "restricted code" when talking with their children rather than the "elaborated code" of middle-class mothers.

Deutsch and his associates (1965) have reported that family interaction data from their studies indicate that "as compared to middle-class homes, there is a paucity of organized family activities in a large number of lower class homes." Like Milner, Deutsch suggests that there is less conversation at meals, as meals are less likely to be regularly scheduled family affairs. Unfortunately this particular paper does not describe in any detail just how these data were collected.

It is probably no accident that nearly every researcher since these reports has included as a part of his enrichment program the teaching of the child to use language effectively. These are typified by the programs of Hess, Caldwell, Gray and Deutsch reported in Hess and Baer (1968). Programs such as those of Bereiter and Englemann (1966) have introduced very formal programs in language training and they base this procedure upon reports that the typical lower class child speaks in incomplete sentences and has such an inadequate syntax for his language that he cannot be understood.

Another interesting contribution to the discussion of this problem comes from the animal literature on deprivation of sensory stimulation. It began with the studies of gentling of young animals as a contribution to their growth, but soon it became apparent that some of the stimulation that had acceleratory effects was far from gentle. Levine (1960), Scott (1960), and Solomon (1964) showed the effect of various sorts of early experiences, harsh, gentle. and innocuous, on animal development.

The logical implications of these studies, if they were actually assumed to be valid for the young human animal, would have suggested all sorts of physical stimulation - visual, electric shock, noise - as means of stimulating development.

J. Mc V. Hunt (1965) is perhaps the most honest in his assessment of this literature, and in his recognition that little of it has been confirmed on human beings or is based upon empirical data about the actual interaction patterns that exist in so-called deprived homes.

He cites the "probable" nature of the deficit from cultural deprivation and suggests that perhaps the overcrowding of the lower class home might lead to a stimulating first year of life, but suggests that the limitation of a linguistic model appropriate to the later school experiences, and the lack of interactions associated with poverty lead to cognitive deficit. He, like many others, suggests that there are a small number of playthings in the homes of deprived children and little room in which to play. He begins his recommendations as to pre-school education to remedy these effects with the phrase "if this armchair analysis has any validity"... again stressing the lack of much empirical information. A review of the literature contained in a recent research proposal (McCaffrey, 1968) reports the same lack of data concerning mother-child verbal interactions in ghetto homes.

At the time we began this study, we generally believed all of these hypotheses, yet we felt that the field was badly in need of empirical documentation of these hypotheses. We wanted to develop measures to describe day-by-day interaction of the mother and the child, and assumed that we would find a striking contrast between the mother-child pairs in a white upper-middle class sample and those from a Harlem lower-class sample. Such a contrast would be valuable for establishing the validity of our methods.

Instead what we have found is a very pervasive pattern of mother-child interaction in which the differences between the upper-middle class family and the Harlem family are only minor variations on a theme rather than completely different tunes, and even these minor modulations are not readily translated into explanations for the Harlem child's difficulty in school or into recommendations for the most effective

type of pre-school programs for children.

Despite all of the statements about the differences between lower and middle class environments, it was apparent even in 1964 that such differences were not well described by the instruments then available. Most researchers were agreed that the only clear way to explore these differences was to obtain a much more complete record of mother-child interaction in various families and thus to obtain descriptions of family interaction that were not merely stated in terms of social class or in terms of amount of interaction.

One cannot deny the facts of real economic poverty in some parts of our society, and of the disgraceful conditions in which many of our poor families are forced to live. There is no denying the fact that children from some of these homes often are not as successful as other children in meeting the requirements of the educational programs of our schools. What is sorely lacking is concrete information about the real deprivations which these children suffer. In some ways they are not deprived of cognitive input despite economic poverty and sub-standard living conditions. It also seems clear that until these cognitive deficits and the strengths are more clearly understood, we are not in a very strong position for recommending appropriate educational programs.

Thus this program of study is based upon several premises about the importance of the study of socialization behavior in naturalistic situations and the problems associated with such an endeavor.

Since socialization and education take place in the home, school, and in peer groups, the first premise is that they must be studied in these naturalistic situations. It is no secret that the naturalistic description of the environments of children in our culture has lagged far behind the experimental investigation of learning processes and personality mechanisms, and has also lagged behind the naturalistic description of child-rearing of other species and of child-rearing in exotic cultures.

A second assumption is that socialization must eventually be described in terms of the actual moment-by-moment interaction of the child and the socializers in his environment, i.e. the relationship of the acts of the socializing agent to the acts of the child. Research involving global variables like warmth, child-centeredness, and the like can make important contributions to knowledge of socialization, but such variables are at best summary statements describing some persistent feature of the moment-by-moment interaction. Because there is no clear behavioral definition of such global variables as warmth, democracy or child-centeredness, it has been difficult to integrate the research on child rearing and education with the theories of learning and social influence that are obviously basic to the processes. Bijou has recently advanced this same argument forcefully (1968). The strategy

of the present research is to describe adult-child interaction in terms of the sequence of acts of the interacting people.

A third assumption of this research is that interpersonal behavior is based on and guided by the actor's intuitive theory of human behavior. When one person tries to influence another person's behavior by exercising his own authority, or by appealing to the other's sense of obligation, or by trying to convince the other that some course of action is in his own best interest, these social actions are all based upon a common sense belief system about the sources of human action (Heider, 1958).

In another research program we are experimentally studying several major aspects of this intuitive theory of human behavior that are relevant to cognitive development and cognitive socialization. The implication of such an intuitive theory for the observation of adult-child interactions is that one can use it to identify socializing actions in a way that describes their psychological meaning to both the adult and the child. When a mother tells her child "You can do it, keep trying" her comment is based on several beliefs: (a) that increased effort will increase the chance of success; (b) that the child's belief in his own ability will encourage him to try hard; (c) and that it is more satisfying and educational for the child to achieve success on his own than for his mother to perform the task for him. These are assumptions in her intuitive theory of child psychology. She may encourage the child in various ways, using different words, and employing different acts, but such acts are psychologically equivalent and should be given a single label. This assumption underlies the development of an interactional language of observation.

The descriptions of socialization are complicated by the fact that children acquire not only their adult-like behavior but also their intuitive theories of human behavior. Very young children do not understand the meaning of adult behavior in the same way that older children do. This sense of the difference between adults and children is also part of our common sense belief system, our intuitive theory if you will, and thus mothers typically treat very young children differently from more mature ones. This is one reason for our emphasis on the study of age differences in adult-child interaction.

The fourth assumption is that the effects of one person's actions on another are not entirely mediated through the meaning that is communicated by the act. There are at least two other aspects of socialization: the subtleties of reinforcement that are often unintended and unrecognized by the socializer, and the effects of modeling and imitation that are only partially intentional. Research in behavioral modification has shown that sometimes the effect of an action is the opposite of its cognitive content. Harris (1967),for example, has shown how a teacher, when she comes over to a child who is sitting alone and urges him to join the group activity, may be reinforcing his unsocial behavior by her attention. She can more effectively influence him by reinforcing him with attention when he does behave socially.

A careful description of the sequence of acts involved in the interaction, together with the identification of acts that can be shown to be reinforcing or not can provide information about the socializing effects of an interaction not carried by its cognitive meaning.

Adults also provide models for the child, sometimes deliberately, in the belief that the child will imitate them, but often quite unintentionally. Physical punishment, for example, which is inflicted in the belief that it will inhibit antisocial behavior, may also provide a model of aggression that the child picks up. The parents' speech throughout early childhood also provides a model from which the child learns his native tongue. This effect of language modeling is such an important part of cognitive socialization that we are not satisfied to record only the specific acts in accordance with their cognitive effect, but feel we must also obtain an actual transcription of the language exchange between the adult and the child.

These then are the premises on which our research is based: (a) the importance of naturalistic observations of the actual moment-by-moment interactions that mediate socialization; (b) the importance of couching these observations in terms that are compatible with the intuitive psychological theories of the people interacting; (c) the necessity for recording the sequence and the contingencies in the interaction in order to search out reinforcement schedules inherent in the interaction, and (d) the importance of an actual record of the language itself in order to assess the effects of the interaction on the development of the child's language.

RESEARCH OBJECTIVES:

The specific objectives of this program of studies were outlined as follows in our original proposal:

1. To devise methods for the collection and analysis of adult-child interactions in naturalistic situations. The data collection is one aspect of the problem. A second aspect is the construction of a behavioral language into which raw observational records can be translated.

2. To describe systematically the total set of interactions that constitutes an environment. This will include some measure of the variety and frequency of various kinds of socialization interactions between adults and children. It is not unlike a sociogram except that the connections between people represent kinds of interpersonal actions, not sociometric choices.

3. To obtain records from a wide variety of settings in order to explore the range of differences that exist in different environments, with children of different ages and different social-economic levels. This material will be used to validate the methods developed in 1 and 2, and to reveal variables for further investigation.

The following chapter will describe some of the methods we have devised or adopted from other studies to describe the concrete information exchange between mother and child, to describe the syntax of the language used by the mother in talking to an adult and to the child in playroom interaction as well as the syntax of the child. In addition we will describe an "interactional language" in which an observer may describe the psychological features of the interactions.

Chapter 2

METHODOLOGY

I. SAMPLE

The total sample for this study consists of three subsamples labeled (1) old longitudinal (2) new longitudinal (3) cross-sectional. The old longitudinal sample consists of 23 mother-child pairs observed at six month intervals, the Negro children (N=12) from age 2-1/2 to 4-1/2, and the middle-class white group (N=11) from 2-1/2 to 3-1/2 years of age. Both groups are being observed again at age 5 but the data is not available for this report. See table 1 for the exact number of subjects in each session.

Table 1
Number of subjects in each sub-group

Sample	Age					
Old longitudinal	2-1/2	3	3-1/2	4	4-1/2	5
West Harlem	12	10	10	11	11	(10)*
Washington Square	11	11	10			(10)*
New longitudinal (West Harlem)				10	10	(10)*
Cross Sectional						
Age 3 (West Harlem)		10				
Age 5 (West Harlem)						20
Total records available	23	31	20	21	21	20

*These records have been or are being collected, but they are not available for this report.

In the fall of 1969, we decided to concentrate our attention on age differences and social class differences within the Harlem community. We therefore increased the number of subjects in our longitudinal sample to 20, 10 of whom were lower class and 10 middle class. We also added a cross-sectional sample of 10 to our three-year-old group in order to compare age and social class differences more reliably. We

also collected data on a group of 20 five-year-olds, half of whom were middle class and half lower. Thus 20 three-year-olds, 20 four-year-olds and 20 five-year-olds are available for comparison. All the children in all the samples are boys.

All these subjects have been selected using the following procedures and criteria.

The Negro mothers in the longitudinal sample are a sub-sample of the children being studied by Dr. Francis Palmer at the Harlem Training Center. All the mothers and children at the beginning of the study lived in West Harlem, where the children had been born in four of the hospitals in the area. Half of the mothers in the Palmer sample are lower class, half middle class. Eleven of the longitudinal mothers remained with us during 3-1/2 years of research; ten of them were present for every session. They range from mothers on ADC to the wife of a Ph.D. trained psychologist. Seven of these families were intact, four had a mother or mother and grandmother as adults in the home. Four had two older or younger siblings, four had one sibling, three were only children. One of the findings emerging from this and other studies is that there is no typical "ghetto family" -- there is a very wide range of education and income within the Harlem community. The Washington Square group of mothers are white and middle class, with husbands in business or professional positions.

II RECRUITMENT OF SUBJECTS *

The cross-sectional samples were recruited independently using the procedures developed by Palmer and his staff at the Harlem Training Center. The names of possible subjects were procured at the Bureau of Vital Statistics in New York City. In this bureau are copies of the birth certificates of all individuals born in the borough of Manhattan since the late nineteenth century arranged according to their arrival date at the bureau.

The birth dates of the three-year-old sample range from 8/65 to 2/66 and for the five-year-olds the birth dates range from 10/63 to 5/64. At the time of the session every child was within a month of his birthday. Since half of each age group was to be middle class and half lower class we collected the names and relevant information for four times the number of required subjects. The volumes for the appropriate birth dates were searched for birth certificates meeting

* This section of the report was prepared by Bonni Seegmiller.

the following criteria:

Borough of residence: 1 (Manhattan)
Area District: 03-38; 85.10 - 85.20 (Harlem)
SCP (Sex, Color, Ordinal Position of Birth): males, Negro, single birth
Father and mother must be Negro
Mother must not be a drug addict
Mother must not be in the last stages of syphilis
Weight: 5-0 lbs. or 2268 g. minimum
With no clear cut mental abnormality such as hydrocephaly or cerebral palsy
Mother's age: 15-45

The above and also the following information was recorded for each potential subject:

Child's name
Date of birth
Father's name, age, birthplace
Father's occupation and kind of business
Mother's maiden name
Mother's age, birthplace
Number of children previously born and now living
Mailing address
Private patient or general service
Position at birth

Following the initial gathering of names of potential subjects, names and addresses were sent to the Customer Relations Office of the General Post Office for address verification and Zip Coding.

Of 308 cards sent to the Post Office, 151 were returned as addressee unknown, or address unknown and were therefore excluded from our sample; 157 were returned with either the old address verified or a forwarding address given.

The next step was to send letters describing the research to all of the families with verified names and addresses. The letter said one of our staff would call personally to discuss the mother's participating in the study. A copy of this letter is in Appendix A. Of these, 15% were returned marked addressee unknown or address unknown. One of our staff visited each of the other potential subjects, explaining more about our study and asking her participation.

A visitor went to the homes of 105 potential subjects from Harlem. In some cases the home could not be located, sometimes no one was home, sometimes the mother or child was not living there and a small number either refused to participate or never kept an appointment. It is important to note that our sample of 50 is biased toward a fairly stable,

highly motivated group of mothers, regardless of social class.

Washington Square Sample

These mothers were located by sending to a play group in the Washington Square area a letter similar to that sent to the Harlem mothers. All lived within walking distance of New York University. All families were intact, and according to the Hollingshead Index belonged to classes I and II. Nine of the children had a younger or older sibling, three were only children.

Social Class Criteria

The Hollingshead TWO FACTOR SCALE OF SOCIAL POSITION based on occupation and education was used. In the preliminary judgments based on birth certificate information, occupation was used to provide a quick index of class. The Hollingshead scale classifies occupations into seven categories with category 1 (higher executives, proprietors of large concerns, and major professionals) the highest, and category 7 (construction workers, maintenancemen and welfare) the lowest.

More complete information was available after talking to the potential subject. At times, the occupation of the husband had changed during the past few years, and also information about education was now available. This factor is also classified into seven categories, from 1 = graduate professional training to 7 = less than seven years of school.

Occupation and education are combined by weighting the individual scores obtained from the scale positions. The weights for each were originally determined by Hollingshead using multiple correlation techniques. The weight for occupation is 7, and for education 4.

The description of the West Harlem sample in terms of occupation and education is shown in Table 2. Those families where the education and the occupation are at the same level are the cells along the central diagonal of the table. Those whose occupational level is higher than their educational level are below and to the right of the diagonal, while those families where the head of the family is underemployed for his educational level are above and to the left of the diagonal. In this sample of 50 families, 33 are under-employed while only 4 have occupations higher than educational level. This confirms the findings of other investigators about the difficulty of Negro men obtaining employment that is compatible with their education. Billingsley (1968) reports that about half of all Negro adults have less than a high school education; in our sample it is 34% because of our efforts to recruit more middle class families. Billingsley also reports on the education, occupation relationship.

Table 2

Number of families from the Harlem sample in each educational and occupational level (N=50)

		IV		III		II			
Ph.D	1					1			I
College Grad	2						3		
Some College	3	2	2	4	5	2			II
H.S. Grad	4	4	4	2	3	1			
Some H.S.	5	5	2	1	1				
Jr. Hi	6	2	2						
< 7 years	7	2	1	1					III
		7	6	5	4	3	2	1	
		V		IV					

I

"... Negroes with similar education to whites do not have similar job opportunities, and negroes with similar jobs do not get similar pay ..." (p.88).

"In fact Negro men must have one to three years of college education in order to equal the earnings of a white man with less than an eighth grade education. After completing college and earning a master's degree the Negro man can count on earning only what a white man can who has graduated from high school." (p.88) In our sample also Negroes with some college education are found in occuaptional levels 3 through 7 with a mean at class 5, which corresponds to an expected educational level of 'some high school'.

Not only are Negroes under employed, but the Hollingshead scale weights occupation higher than education. The result is that by the Hollingshead criteria, half of our sample is in class V, the lowest social class in his scale. This is the dividing line between lower and middle class in this study. In terms of social prestige, both in the eyes of whites and blacks, this is probably accurate. Billingsley again reports that probably 50% of Negroes consider themselves and are considered by others to be in the lower class, composed of the working non-poor, (semi skilled job holders) the working poor, (largely unskilled laborers) and the non-working poor, (unemployed and probably on welfare).

For the analysis of mother-child interactions we have felt that perhaps education is a more significant variable than occupation or social prestige. Therefore in several analyses we have compared groups from different educational levels, rather than solely from different social classes.

Experimental Situation

We have already stated our belief that important cognitive development and socialization takes place in the child's day-to-day interactions with his environment. The playroom was designed therefore to stimulate the child's curiosity and information-seeking activity. It contained a variety of activities: inside-outside jigsaw puzzles; magnet board with letters, people, and cars; a lock box; balance beam; wooden train set; barn with animals; and a doll house with dolls set up in a standard way for each session. Materials were always arranged in a standard way, as shown in figure 1.

When each mother and child arrived at the laboratory the experimenters talked with them informally for a time, and when they seemed relatively comfortable the mother and child were taken to the playroom.

The instructions to the mothers were:

"You know that we are studying the ways in which young children learn about their world. Part of this is through talking and playing with toys and with other people. We know (David)is probably more

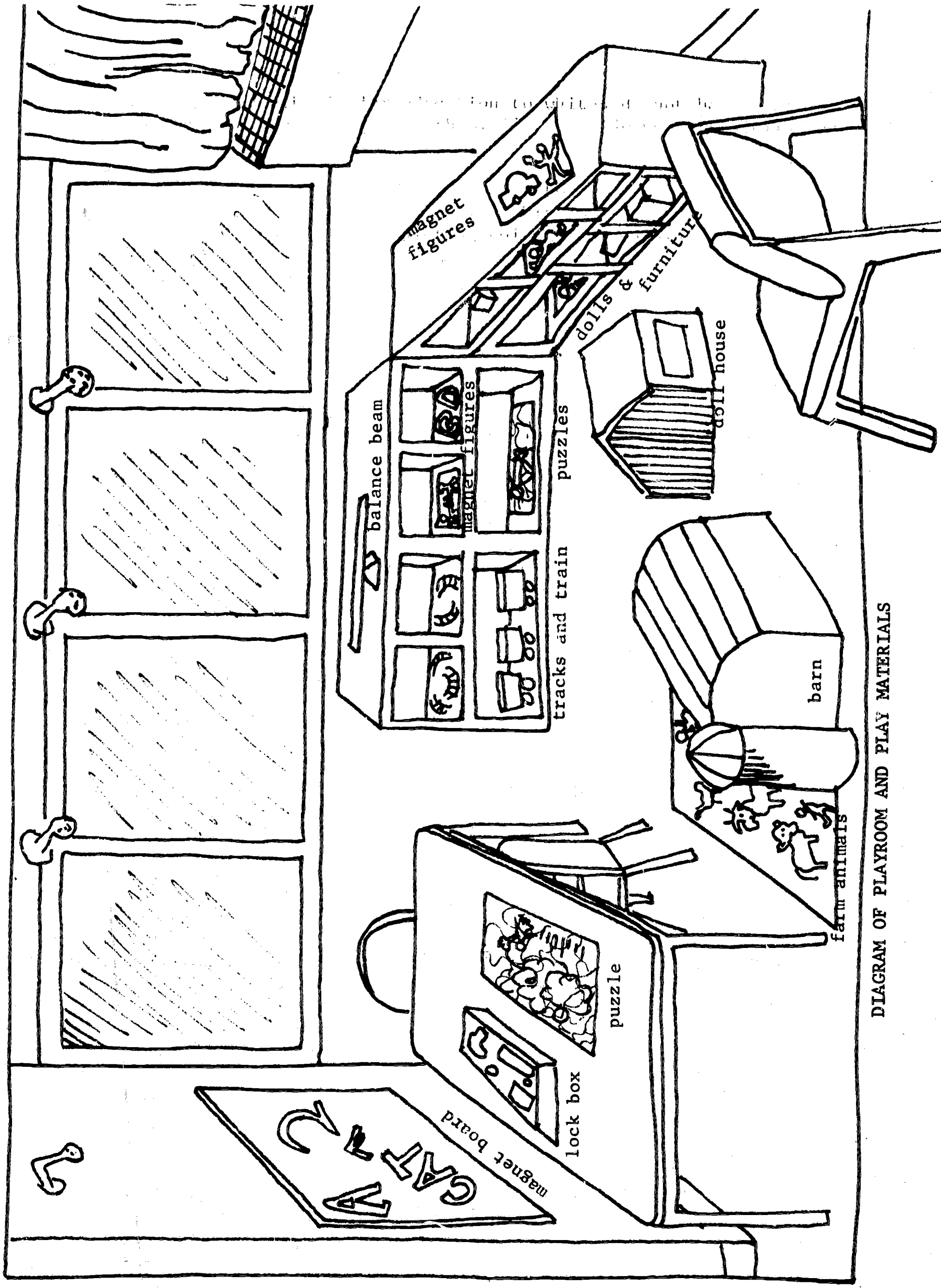


DIAGRAM OF PLAYROOM AND PLAY MATERIALS

comfortable with you than anyone else, so we'd like you and (David) to play here together for a half an hour. You can play with him in any way you like. We know it won't be quite the same as if you were home, but it would be helpful to us if you could talk with him about the same way as you do there. Then Mrs. S. will take (David)to do some other things, and I would like to talk with you a bit."

The mother and child were then left alone for 30 minutes in the play room. The interaction was recorded on audio tape, a modified running record in interactional language was dictated by an observer into a tape recorder, and the child's non-verbal exploratory behavior was recorded by a pre-coded category system.

The play session was followed by an interview with the mother concerning people with whom the child interacted, play materials, and her assessment of the naturalness of his play here. During this time language measures were administered to the child. Copies of these measures can be found in Appendix F.

At the end of each session, all observers were asked to rate the session on a number of variables. A copy of this rating scale is also in Appendix F.

A review of the literature four years ago, when this study began, revealed few measures for studying the cognitive aspects of mother-child interactions; therefore one of our major tasks was to identify and label specific behaviors important to the study of information-seeking, and to develop techniques for measuring and analyzing these behaviors. We have developed instruments for measuring both the non-verbal and verbal acts of mothers and children. A description of these various measures follows in the next section.

IV. METHODS

VERBAL INFORMATION EXCHANGE (VINEX)

This category system was developed to measure the quantity and quality of the verbal exchange between mothers and their pre-school children. Data were collected in the semi-structured play situation just described. The raw data consist of transcripts of an audio recording of the actual conversation which took place in this playroom. An observer simultaneously described the nonverbal and verbal behavior he thought to be important for the cognitive development of the child. Both these records were used to code the audio transcript. An example of the completed record can be found on page 22 .

VINEX

Utterance

The unit of the verbalizations is the utterance. The utterance is basically a sentence, though not necessarily a complete one. "A burn up

place", "Oh", "Let's make _____", are coded as utterances as well as "I can't stay in this school." Utterances are identified by the coders by intonation and pauses. A drop in pitch seeming to indicate the end of a thought, a question intonation, and a definable pause between utterances were cues to the transcriber that an utterance was ending. Often an utterance in these transcripts is followed by an utterance of another person, though there are many examples where either a mother's or a child's monologue is several utterances long.

The division of the record into utterances represents the best pooled judgement of a transcriber and an observer, both of whom were present in the observation booth during the play session. The transcriber took shorthand notes during the session in order to help her better transcribe the sessions.

Information Coding

Each utterance was coded for its form - was it a question, statement or incomplete. Information content was coded in terms of its quality - was it permanent, transient, behavioral or fantasy. The mode of the information, and the response qualities of the utterance were also coded. Definitions and examples of these major categories and their sub-categories are given below.

VINEX Category Descriptions

Form — Each remark must be classified according to one of the following form categories.

QQ	Questions which request information. All "wh" questions. "What is this?"
QH	Questions which present a hypothesis and can be answered by yes or no. "Is this a blue car?"
QA	Questions which request attention. "See this car?"
QR	Questions which request behavior. "Will you get that for me?"
SS	Statements which provide information that the speaker is certain of. "This is a car."
SH	Statements which present or comment on a hypothesis. "I think this is a dog." Speaker is uncertain, doubtful.
SA	Statements which seek to elicit attention. "Look at this."
SR	Statements which seek to elicit behavior. "Get that piece of track."

OY — Utterances which cannot be coded, are incomplete or incomprehensible. "Crying, laughing, noises, ..."

OMIT — Utterances containing blanks in transcription. "Give me the _____."

OC — Repetition of own statement for sound. "This is a seat. Seat, seat, seat...."

Information Content

Each utterance, except those already coded as OY, OC or OMIT, must then be coded according to one of the following information categories.

P — Permanent information: Utterances involving a general principle or property which has enduring qualities, i.e., where the information, no matter how modest, can be carried over to the future. These may refer to the environment, both immediate and remote (including toys), or to repeatable procedures and states. "What is this thing for?" "This car is bigger than that one." "Cows say moo." "Babies are happy when they've been fed."

T — Transient information: Remarks referring to real passing events where the information has no future applicability. Those directly concerned with fantasy play are not included in this category, but reality based utterances occurring within a fantasy context are included. "Where's the ball?" "Do you want a cookie?" "May I play with this now?" "Put that over there." "Please help me fix this." "Let's play house."

F — Fantasy information: Remarks directly involved in "pretend" play where the child may know that the information is not reality based. "I'm the mother, you're the baby." "Does this train go to Africa?" "Now it is nighttime so we go to sleep?" Also acts performed by the animals or dolls, rather than the child himself. e.g. "Mommy is cooking a hamburg."

B — Behavioral information: remarks which carry no other information than comments on the actions of the people involved or the persons own actions. "Come over here.""Get off the table." "I am doing this."

A — Approval: approval is expressed without other information being conveyed. "You are doing really well." "I like the way you are fixing that."

D — Disapproval: disapproval expressed without other information being conveyed. "I don't like you to do that.""Naughty boy."

Mode of Information Exchange

a	Explains:	Gives or requests information about causality, function, purpose, goal or intention. "Why is it raining?" "You can't do it because it's dangerous." "What is that button for?"
b	Limits:	Gives or requests information about the possibilities or constraints inherent in a situation. "You may not scream here." "This is too large for the space." "I'm not big enough to do this yet." "The track can't go there."
c	Clarifies:	Gives or requests further information or repetition. "What did you say?" "This is the one I mean."
d	Describes:	Gives or requests a physical description, representations of actions, or appearance. "Cat's eyes shine in the dark." "Is this bigger than that?"
f	Feeling:	Gives or requests a description of feelings, or wants "I want to go home." "I am tired."
g	Demonstrates:	Shows how to do something while talking about it. "It goes into the puzzle this way." "Show me how to do it."
k	Commands:	Requests behavior with insistence or anger. "Get down right now!" "You must do it immediately."
l	Labels:	Gives or requests the name of an object or person No other information given except name. "This is a dog." "What is this?" This includes the labelling of letters within a word.
s	Specifies:	Requests or provides information about location and selection. Usually can be conceived of as accompanied by pointing. "I will get this one." "Go over there." "Which one shall I do?"
t	Thematic:	Sounds and remarks made as part of the "script" in fantasy. "All aboard." "Moo."
r	Reiteration:	Repetition by same speaker, provides no new information. Not to be confused with "c" which is used when the responding person has asked for repetition or further information about what was previously said. "M: What is that?" "M: What is it, Anthony?"

Response Demand Code:

Code the utterance as to whether it demands a response or not.
Response required +
No response required -

Response Categories:

Every utterance must have a code in the response category column, either coding the response or explaining lack of response.

Categories describing reason for lack of response

OY — Response unclear

OO — Ignores

OM — Use this code for an utterance that follows an utterance which has been coded OMIT or OY.
c:_____ (OMIT)
M: What did you say? (QHTc + OM)

C-C — Use this code when the speaker is continuing talking and
(a) the other person does not have time to answer or,
(b) a response is not required.
M: Do you like the airplane?
M: Hmm? (this is said before the child has time to respond) C-C

X-X — Use this code for the first utterance in a new exchange. An exchange is a series of stimuli and responses about one thing.
M: You spell my name with a "I."
C: (sound)
C: Let's make...
M: Hmm?
C: Spell your name.
M: Come on. Z
What ya looking at? (X-X) New exchange begins
C: Watching burn up place.
M: Watching what place?
C: A burn up place.
M: They must of had a fire.

Z — Code next to the last utterance of the previous topic. Change of topic is generally when referent changes. Ex. "puzzles" to "lock box."

Direct

A direct response exactly answers a previous utterance. It need not necessarily be a response to the immediately preceeding utterance. It can be a response delayed by interjected utterances, but it must be

one that is a direct answer to some previous utterance.

M: (asked question)
C: (initially ignores)
C: (talks about something else)
C: (remembers mother's question and answers it directly)

Use Prefix "1" for direct responses.

-0	Misinforms	Gives incorrect information "It's a horse" when should say "It's a cow."
-1	Informs	Gives factual information. "What is that?" "It's a ball."
-2	Confirms	Simple agreement in response to a question or statement. "This is green." "Yes it is."
-3	Denies	Declares remark to be untrue without correction. "Today is Thursday." "No it's not."
-4	Corrects	Provides correct information. "This is blue." "No, it's purple."
-5	Accepts, complies	Accedes to a request or accepts information or opinion. "Okay, I will do it for you." "That's alright with me."
-6	Rejects	Non-acceptance or non-compliance. "No, I won't help." "I don't want that one."
-7	Praises	"That's beautiful. You did it just right."
-8	Punishes	Indicates disapproval or dissatisfaction. "That's very naughty."
-9	Uncertainty	Indicates doubt about information provided or in general. "I'm not sure." "I think it might be blue."

Peripheral

A peripheral response does not exactly answer a question, or responds indirectly to a statement. Prefix "2" is used with some suffixes described above.

C: What is that?
M: Ask your father. (2-1)

C: What is in those boxes?
M: Things for other people. (2-1)

No information

3-1	Direct question in response to the preceeding utterance. "What did you say?" "What do you mean?" "I should do what?" "How can I do that?"

3-2 Peripheral question in response.
M: This is a balloon.
C: Where did you buy it?

4-0 Remark reflected back.
C: What do you think it is?
M: You tell me.

4-1 Encouragement. "Try again." "Keep on pushing." "Go ahead."

5-0 Parroting.

8-0 "Don't know" responses.

9-0 Delayed. "I'll tell you later," incomplete, or answers self immediately.

3-3 Repetition in response to an ignoral.
M: What is that?
C: (0-0)
M: What is it, Anthony? (3-3)

Table 3 is an excerpt from the record of one of our 3-1/2 year-old Harlem boys. It includes the narrator's record in interactional language, the verbal transcript of the mother and child, and the VINEX coding for each utterance.

Reliability of VINEX Coding

Several different people have been involved in coding the transcripts of the verbal interaction in terms of the VINEX category system. Midway in the project, because of a move of the principal investigators from New York University to Cornell University, the entire staff of coders was changed at one time. These factors pose problems for the maintenance of reliability. Reliability checks were made among the group of coders who did much of the coding before September, 1968. and each new coder was given practice on already coded transcripts until he showed an acceptable reliability. When the entire staff was changed, the Cornell staff coded transcripts that had already been coded by the New York staff; problems of reliability were discussed until finally the new staff achieved acceptable reliability levels both with previously coded records, and within the new staff itself.

The agreement on the coding of the form is on the average 90%, with individual pairs of coders ranging from 79% to 94%. Two-thirds of the pairs have agreements above 90%. On coding the content of the utterances, the mean percentage agreement is 84%, with individual pairs ranging from 72% to 91%. Three-quarters of the pairs are above 80%. The reliability of coding the mode is lowest, with a mean of 70% and a range

TABLE 3

SAMPLE OF MOTHER-CHILD INTERACTION
COMPLETED RECORD AND CODING

Subject No.
Age 3-1/2 Session 3

Minute	Sent. No.	Narrator's Record in Interactional Language	Utterance Number	Verbal Record Mother & Child Utterance	Sent. No.	Subj. Code	Alph. Code	Resp. Quality
15	224	He picks out more letters.	287	C: Inez we clean up, we gotta go home, right?	287	C	QHBd	
	225	Sticks them back to back.	288	M: Right.	288	M	SSBd	1-2
	226	He's talking about spelling his name.	289	C: I can't stay in this school.	289	C	SSTb	
	227	Says he spells with an "I"	290	You gotta tell school spell my name with an "I".	290	C	SRP1	
	228	She, she says he spells it with a "K".	291	M: No you spell your name with a "K".	291	M	SSP1	1-4
	229	He's struggling with the junction box.	292	C: oh.	292	C	OY	1-5
			293	M: You spell my name with a "I".	293	M	SSP1	
			294	C: (sound)	294	C	OY	
			295	C: Let's make ...	295	C	OY	
			296	M: Hmmm?	296	M	QQTc	3-1
			297	C: Spell your name.	297	C	SRP1	2-1
			298	M: Come on.	298	M	SRBd	4-1
			299	What ya looking at?	299	M	QQTd	
	230	He requests some kind of a definition of a place.	300	C: Watching burn up place.	300	C	SSTd	1-1
	231	She seems a little annoyed, also today	301	M: Watching what place?	301	M	QQTc	3-1
			302	C: A burn up place.	302	C	SSTC	1-1
			303	M: They must of had a fire.	303	M	SHTd	
			304	C: They did.	304	C	SSTd	1-2
			305	Still there.	305	C	SSTd	

Minute	Sent. No.	Narrator's	Utterance Number	Verbal Record Mother & Child Utterance	Sent. No.	Subj. Code	Alph. Code	Response Quality
	232	He takes out the street puzzle.	306	M: Why don't you put the puzzle together.	306	M	QRBd	
	233	Replaces it in the bookcase	307	C: I don't wanna do that.	307	C	SSTf	1-6
			308	M: Don't you have to put the A.B.C. back?	308	M	QHTb	
	234	He replaces the alphabet letters in the box.	309	C: O.K.	309	C	SHTd	1-5
16	235	Returns them to the bookcase.						
	236	She directs him which shelf to put it on.	310	M: Put it back where you got it from.	310	M	SRBs	
	237	He complies.						
	238	He returns to the puzzle, street puzzle.						
	239	She suggests that he do the airplane puzzle.	311	You don't want to fix the plane?	311	M	QHTf	
			312	C: Huh?	312	C	QQRx	3-1
			313	M: See the plane	313	M:	QAP1	3-2
	240	He complies.	314	C: Yeah I wanna do it.	314	C	SSTf	1-2
	241	He asks if he should do it by himself.	315	You know fix this by myself.	315	C	SSBd	
			316	M: You gonna fix it by yourself now?	316	M	QHTc	
	242	He dumps the pieces out on the table.	317	C: Yeah.	317	C	SHTc	1-2
	243	He draws her attention to the figures on the back.	318	C: Look.	318	C	SABs	

from 59% to 75%. It turned out that nearly half of the total disagreement between coders was a failure to discriminate between describes and specifies. For most analyses therefore we collapsed these two categories.

For the coding of response quality, it is necessary first to identify an utterance as a response, and then to code its features. The reliability of both of these judgments is generally above 80% with isolated pairs of raters going as low as 70% on occasion.

SYNTACTIC COMPLEXITY

The taped record of the mother and child's verbal interaction was used for two purposes. One was the VINEX Coding as described in the last section. The taped record is also the raw-data for the assessment of the syntactic complexity of both the child's and the mother's language. The syntax was studied both quantitatively and descriptively. Several indices estimate the general level of syntactic complexity. In addition some 89 descriptive features of syntax are marked, e.g. yes-no questions, subordination of clauses, conjoining of clauses without conjunction, arrangement of nouns in a series, etc.

From these data we can examine such questions as the difference in the mother's language when interacting with her child and when talking to an adult interviewer; the differences between the mother's and the child's complexity when interacting; and the changes in complexity occurring with age. From the description of the syntax we can look for the age of acquisition of some specific grammatical features and to what degree the mother's speech is a model or a preview of what the child will acquire.

The measure of complexity was developed by Dr. Sheldon Frank and Dr. Harry Osser in their study of language development in Negro children in Baltimore. This method was chosen (1) because it is applicable in these naturalistic situations, (2) because it is a differentiated measure that makes relatively fine discriminations among sentences, (3) because it depends on many elements in the sentence and so should reflect many potential aspects of language development, (4) because it seems to be closely related to intuitive feelings about the complexity of a sentence, and (5) finally because Dr. Frank was available as a consultant to our project. He trained the coders, supervised the scoring of the records, and has collaborated fully with us in the analysis of the results.

The measure itself is based on the concepts of generative grammar although like many psycholinguistic studies the syntactic theory on which it is based has become outmoded by the time the study is done.

Chomsky as of 1957 and more specifically Roberts (1964) describes sentences either as kernel sentences or as more complex ones, derived from kernel sentences by transformations. Frank and Osser's is based upon the idea that the kernel sentence consists of a nucleus —the bare

bones of the noun phrase and verb phrase -- plus possibly optional additional element like an adverbial phrase. Then the kernel sentence is transformed into the final output through the application of one or more transformations. Each transformation involves one or more of the following elements: addition of an element, transposition of an element from one place in the sentence to another, the deletion of some element, or the addition of supersegmental elements like an intonational change. The basic assumption for the scoring is that each of these changes in a sentence, whether in the kernel or through transformations, is of equal difficulty. Therefore the score is the number os such additions, deletions, transpositions, and intonational changes.

A complete nucleus like "we played" receives a score of 2, one for the noun phrase, one for the verb phrase. Some sentences that occur in conversation may lack one of these elements, e.g. "Huh?", or "A car," in answer to the question "What is that?". These utterances receive scores of 1, and there are many in our records.

"I never played only one time" is scored 5. The nucleus "I played" receives 2. The additional optional elements are "never", "one time", and "only". Each adds 1 to the score.

"We sing songs and play." is scored 4 -- 2 for the nucleus and 2 for the transformation that adds a second sentence to the first with an "and". If the second sentence had been complete, "We sing songs and we play", the score would have been 5. One might argue that the original form is obtained from the latter by a deletion of the noun phrase from the second sentence and therefore ought to receive an additional score. Here Frank departs from the strict logic of transformational grammar and argues that the many deletions that occur in generative grammar do not count as additional complexity. This is certainly a debatable question, but the important thing for the present study is that the scorer need not debate it. Each transformation has an assigned score so that all instances of it are scored the same way.

Here are some more examples of sentences and their scores.

"We sit on the circle and stand on the circle."	Score 6
"What is that?"	Score 5
"You set them on the floor because they can't stand on the rug."	Score 12
"Well, you go over and see what it is."	Score 16
"Ain't that the thing that say, oink, oink, oink?."	Score 11
"Yea, stand it up and see if you know what it is."	Score 21

The analysis can be made in various ways to study the appearance of particular transformations and specific optional elements.

In all there are 89 different transformations that have been assigned scores, so that the actual coding of a sentence consists of indicating whether the nucleus was complete, what optional elements were included in the kernel, what transformations were applied and what optional elements were added to the transformed sentence.[1] The mean score per sentence will be referred to as mean complexity.

There is considerable debate whether the complexity involved in the production of a sentence is reflected by the number of transformations that would be required to generate this sentence within generative grammar. Nobody now believes that verbal production is a direct reflection of grammatical generation. Furthermore there is no question but that the present measure is as much a measure of the semantic complexity of the sentence as its syntactic complexity. Many high scores reflect the addition of more and more elements of meaning into the utterance as well as the rearrangement of these elements in accordance with syntactic rules. What emerges, however, is a score that reflects many of our intuitive notions of complexity. It ranges from 1 as a lower limit but it has no theoretical ceiling. Sentences scoring above 20 are quite rare.

More important, the appearance of grammatical features like question intonation, yes-no questions, what questions, affirmative tags etc. are identified individually. It is valuable to see when and in what order they appear and how they are related in the child's and mother's speech, regardless of whether they are best described as transformations of a kernel sentence or as descriptive features of the language. The concepts of transformational grammar have certainly guided the analysis, but much of the value of the analysis is not dependent upon the validity of transformational grammar as a production model of the child's utterances.

---

1 A coding manual for scoring utterances according to this system is presented in Appendix A . While it may not be necessary for the child to know transformational grammar to utter the sentences, it is essential for the prospective coder to be familiar with the general notation of transformational grammar to code utterances or even to understand the coding manual. In training coders in this project, Frank found Roberts (1964) very useful. Frank's modifications of what Roberts presents can then be understood from the coding manual. It took several weeks to train relatively sophisticated psychology undergraduate and graduate students to code these utterances. Training can be shorter for someone already familiar with transformational grammar.

NON-VERBAL EXPLORATORY BEHAVIOR CODING

In addition to the analysis of the verbatim verbal interaction, two other records of the interaction were obtained; one was a coding of the child's non-verbal explanatory behavior described in this section and a second was a narrative record to be described in the following section. Each non-verbal act by the mother and the child was entered on the coding sheet (See Table 4 for a sample data sheet). One sheet recorded the acts of a one minute period.

The data sheet provides for the identification of the actor, the act, and the object of the action. As long as a sequence of behavior has the same actor, action, and object, the same act is considered to continue, even though several motor movements might be distinguishable. The coder records a new action whenever the actor, the action or the object changes.

The category system is intended to describe the non-verbal behaviors that are not recorded on the tape of the verbal interaction, but when the mother and child were just talking the coder felt uncomfortable to be writing down nothing, so we allowed her to record talking and questioning as actions provided that no non-verbal behavior was occurring. The central task was, however, to record non-verbal behavior.

The coding categories indicated on the data sheet are described below. The categories are arranged in four main groups, gross motor, visual, fine motor and miscellaneous. Within each main group the categories are intended to discriminate between aimless behavior and gaol-directed behavior. The goal-directed actions are further subdivided into categories that reflect such cognitive functions as degree of attention required, information seeking behavior, and far-sighted goal-directed behavior. For analysis the visual, goal-directed motor, and non-goal-directed motor categories were combined. Thus several descriptive indices can be constructed to highlight different aspects of exploratory behavior.

Coding Categories:

1. Gross Motor

GM — Gross motor - jumps, runs, dances, claps, bounces, sits,etc.

W — Wanders - moves aimlessly around the room

APP — Approaches - physically moves towards an object.

Table 4

Sample Mother-Child Exploratory Behavior Coding Sheet

Name: John Doe Date: 12/5/68 Observer: P.S.D.

Minute: 1

	Verbal				Non Verbal Exploration																					
Act	Sub		Verb		Gross Motor			Visual				Fine Motor						Miscellaneous						Object		
	G	M	T	Q	GM	Wan	App	G	L	S	Ob	Han	Man	Exp	Prep	C	P	Dem	Imit	Pts	Con	Fan	Agg	C	M	Object
1		✓											✓													Doll
2	✓								✓		✓														✓	
3	✓																									Doll
4	✓						✓																			Doll
5	✓	✓	✓																							Doll
6		✓																								
7	✓						✓																			Book Case
8	✓												✓													Puzzle
9	✓														✓											Puzzle
10		✓									✓													✓		

2. Visual

G — Glances - briefly gazes at an object, person, or part of the room.

L — Looks - visually examines by intent looking for several seconds.

S — Searches - intently looks for and moves about to find a specific object or part of an object - e.g. a puzzle piece, the right sized nut, a letter missing from a name,etc.

Ob — Observes - studies for more than a few seconds a person's behavior or the nature of an object.

3. Fine Motor

Han — Handles - handles objects in a random, non-purposeful, and/or casual way - for example, dumping puzzle pieces, pushing toys away, fiddling with a lock.

Man — Manipulates - thoughtfully handles an object, includes physical exploring of an object, as well as purposeful manipulation, for example placing letters onto the magnet board, making a pile of wooden objects, bending a doll's legs for sitting.

Exp — Experiments - physical hypothesis-testing, seemingly with an idea of what ought to result, whether or not the experiment is successful. For example, puts weights on the balance beam and seeing if they balance, tries a puzzles piece in an oddly shaped spot.

C — Constructs - engages in a relatively sustained activity in which the person has a goal and is reasonably competent. For example, putting together track pieces, writing a word with letters, placing one piece after another correctly into the puzzle frame.

P — Plays - a series of motor acts which the child or adult seems to have assimilated so that they are easy for him and form part of an ongoing activity. For example, running the train along the track, placing all the animals into the barn, or dolls into the dollhouse.

Prep — Prepares - begins or finishes an activity by setting up or putting away objects - e.g. turning over puzzle pieces prior to doing the puzzle, placing all the tracks on the floor before making a track, putting away one puzzle so that another one can be done, etc.

4. Miscellaneous

Dem — Demonstrates - illustrates by example how to do something.

Imit — Imitates - repeats an action or a verbal utterance in an imitative fashion.

Pts — Points - physically points to a person or object.

Con — Contacts - approaches, touches, shows or gives something to another person.

Fan — Fantasizes - non-verbally pretends to be someone or something other than oneself, or to carry out an action via an object or person other than oneself. For example, making a horse kick a doll, crawling on the floor as a dog, pretending to eat a letter, etc.

Agg — Aggresses - hits, attacks, destroys, or attempts to destroy an object or person.

Reliability

Two observers coded preliminary sessions simultaneously and discussed disagreements until independent coding reached 70% agreement. Reliabilities for agent and object were above 95%. Most of the disagreements occurred among handles, manipulates, and experiments. The amount of involvement often changes as the child continues to play with a particular object and it is often difficult to judge when the aimless activity changes to goal-directed behavior.

As with all systems for recording naturalistic behavior on the spot, it is virtually impossible for observers to score 100% of the acts, and a major problem is that observers do not always record precisely the same acts in any minute. Therefore the inter-coder reliability is calculated on the acts simultaneously scored by the two observers.

INTERACTIONAL LANGUAGE

An observer was present at the mother-child play sessions who dictated into a tape recorder a narrative description of the interaction of the mother and the child. One of the original aims of this research project was to develop a strategy to allow an observer to describe the interactions of the mother and child in ordinary English and at the same time permit the record to be rapidly and efficiently analyzed by a computer program without having to be read over and coded by still another person. We are just now approaching the realization of this aim.

Interactional language consists of three elements:

1. A vocabulary to help the observer describe human interaction easily and unambiguously without imposing any aribitrary constraints beyond those of trying to be precise.

2. A program for transforming the sentences used by the observer into a standard canonical form which is suitable for analysis. In this step contextual clues are utilized to supply the referents for pronouns and other words whose antecedents are in earlier sentences, and to supply the specific meaning of words which would be ambiguous out of context. e.g. to label "tell" as either a request or as information-giving depending on the context. The sentence in its canonical form has the actor, act, recipient, content and modifier words explicitly labeled.

3. A program for analysis that depends upon assigning a set of semantic features to every word in the vocabulary. The program for the content analysis looks for the presence or absence of various semantic features and counts the gross frequencies and contingent-frequencies among various semantic features.

Each of these elements will be more fully described later but first, why is the observer's record valuable enough to justify the expenditure of so much energy in trying to develop an efficient automated analysis of it?

The study of human interaction involves three steps, the recording of the interaction, the coding of it, and the analysis of the coded material. A video tape of an interaction is a record of it. When this tape is viewed over and over again by coders who then assign the behaviors to a category system, it is coded. Then frequency and correlational analysis may be applied to the coded record. When an on-the-spot category system is used by an observer, the recording stage and the coding stage are collapsed into one. When the observer records his observations in ordinary language, the recording and the coding is again collapsed, but the code employed is a complex one that permits many subtle distinctions but also can introduce ambiguities.

Actually it is impossible to obtain a complete record that contains everything in the interaction. Two video tapes, one of mother and one of the child are probably closest, but the camera man picks and chooses what to focus on, the child's hands working a puzzle or his face. Or if a fixed camera is used, then some actions are out of the field or are too tiny to be observable. The record is incomplete but the loss of information is not intentionally selective.

The investigator, depending upon his problem, the variables he wants to record, his theory of behavior, whether he is exploring a problem or testing a hypothesis, etc., selects some set of procedures for recording, coding and analyzing the data. When the investigator needs to discern or to record many variables whose distinctions are difficult to make and hence potentially unreliable, he needs as

complete a record as possible, which is then minutely coded by human coders after many replays. When only a few variables are needed, direct on-the-spot coding is quick and efficient. It is important however that these decisions be made by the investigator in terms of his problem, not in terms of technical constraints. Some studies have been limited to just a few variables only because the labor of coding records has been too great.

The role of the natural language observation is important for certain classes of problems because the observer using ordinary English has at his command a far larger category system than any observer could be expected to learn de novo. Furthermore, the natural language contains many words that are concerned with interpersonal interaction, like commands, asks, helps, begs, thanks, repays and retaliates. Language has become a medium of communication between people who frequently talk about social interaction. The natural language is not well adapted to describing the phonetics of speech or the motor movements of swimming. It is, however, a culturally developed coding system for molar behavior. Whether the interactional variables inherent in natural language are the ones that are scientifically the most valuable is a debatable question, but they are the variables that other people in a society use in their descriptions, their evaluations, and their responses to an individual.

To put it in other words, the natural language is the notation system that has accompanied the development of intuitive psychological theories of human behavior and social interaction. Fritz Heider has described this intuitive theory under the label, "Naive Psychology" (1958). Baldwin (1967) extended this description to a naive theory of child behavior and socialization. The Baldwins (1969, 1970) have empirically investigated some aspects of naive psychology and have verified the accuracy of certain models underlying common sense judgments of human behavior.

The language stems from naive psychology. Heider, for example, hypothesizes that one of the basic distinctions is between "can" and "try" - which are themselves words in ordinary language. In order to succeed, a person must be able to succeed and also try to succeed. We have observed how many of the interactions of mothers with their children make sense if one distinguishes between the mother's actions which are intended to help him succeed (hints, helps, assists, instructs, informs, etc.) and those which are intended to make him try (encourages, urges, reassures, comforts). We believe that the mother's actions are guided, partly at least, by this intuitive theory about how children behave and what they need to succeed.

The natural language contains many words that highlight the distinctions of naive psychology, and thus is well adapted for describing interpersonal interaction.

At the same time ordinary language is not a perfect coding system. Some frequently recurring actions do not correspond to any single word in ordinary language and in other cases the words of ordinary language are not precise and unambiguous enough. For example a mother frequently helps her child by partially performing the task he is attempting but leaving it to him to complete it. A mother may put a jig-saw puzzle-piece in the correct area of the puzzle but leave the child to fit it into place. We have therefore defined the verb "approximates" to specify this particular kind of help.

To give another example, we have observed at least nine different ways that the mother may respond to a misstatement of a child. Suppose the child labels a calf a "cow". The mother might say:

Mother's statement	Interactional description
No, that's not a cow.	negates
That a calf.	corrects
No, that's a calf.	negates and corrects
Are you sure that's a cow?	doubts
Isn't there another word for it?	doubts and requests child to relabel
Isn't that a calf?	doubts and corrects
That's right it's a baby cow.	confirms
That's right, it's a baby cow, we call it a calf.	confirms and corrects
----------	ignores

We cannot be sure that all these distinctions have different psychological impacts on the child, nor that they necessarily reflect different intentions on the part of the mother, but we suspect they do. We believe that the mother who consistently "confirms and corrects" rather than "negates and corrects" probably produces different cognitive results in the child.

One of the features of Interactional Language, therefore, is that it supplements ordinary language by providing the observer with terms to make precise distinctions without requiring him to learn a whole new vocabulary. In fact the observer need not use the term "negates"; he may describe the mother's sentence by saying, "Mother says no". The computer program can paraphrase this sentence into "mother negates". On the other hand the observer probably will find the term "approximates" easier to use than some circumlucation like "mother puts puzzle piece in correct area of puzzle and leaves child to fit it into place." Not only will such a verb be more usable, but also it will encourage the observer to use the same term in other types of activity, as when the mother trying to help the child remember the color black said "Bl...a ....." and let the child finish the word.

Another problem of ordinary English as used by the observer is that its meaning may be precise enough, but it depends upon complex contextual

cues. For example: The child fits the puzzle piece incorrectly.
The mother removes it and asks the child to try again.
The second sentence is not by itself unambiguous because the antecedent of "it" is in the previous sentence. The human reader knows exactly what is meant but the sentence itself does not say so. This is a problem to be solved by the computer program, and within the relatively restricted syntax of the sentences used by observers we can instruct the computer how to find the appropriate antecedent. We need not demand that the observer never use pronouns with antecedents in an earlier sentence. Perhaps the program will not work perfectly, but if it works 85% or 90% of the time the loss is not unbearable given the large body of sentences that is available for analysis (500 to 1000 for a half hour interaction).

To summarize:

1. We want to capitalize upon the observer's skill as a psychological instrument and provide him with a language that permits him to say what he wants to.

2. We want to supplement and refine his vocabulary of ordinary English so that he can express precisely what he wants without having to search for words or for complicated circumlocutions, and also to help him recognize frequently recurring actions under some generic label.

3. We do not want to restrict his grammar any more than is absolutely necessary. We belive that the grammar used in such observations is naturally restricted enough that we can write computer programs to put his sentences into a canonical form. We have never had an observer who was inclined to use such a sentence as "It is the puzzle piece that the mother wants to help the child place by putting it into approximately the correct spot". These are the sentences the linguist finds fascinating to try to decode, but observers do not use them so we do not need to solve that problem.

Description of Interactional Language

1. The vocabulary

The development of a vocabulary for interactional language is not to write a restricted lexicon which the observer must use exclusively, but rather to provide the observer with a supply of words to make his observations easier by reducing the time necessary to search for a word or to devise a circumlocution. By indicating the kinds of distinctions in which we are interested, the vocabulary also sensitizes the observer to those distinctions, and helps him speak more precisely. But from the observer's point of view it is intended to be an aid to clear expression rather than an imposition of restrictions.

The dictionary does sometimes limit the meaning of words. For example, we would like the observer to use the term "label" when talking about the standard name for an object and to use the term "name" when

a doll or an animal is given an arbitrary name, like "Mary" or "Fido", that cannot be called correct or not. In these cases we have tried to be in conformity with distinctions that generally hold in English usage but perhaps are not used consistently. We use "compare" to describe the indication of similarities between two objects, and "contrast" for the indication of differences. In ordinary language this is not a rigid rule, but our distinction is in general accord with usage.

If the observer does not follow these suggested meanings, the record loses some precision but does not lose its value. The analysis of the record is based upon the appearance of words that have certain semantic features. Thus both "compare" and "contrast" have many common features; they are verbal acts, they give information, they relate two objects to each other. They are different in that one discriminates and one indicates similarity. But only for those analyses where we were specifically interested in that possible difference in verbal behavior would the misuse of the term lead to any error in interpretation.

Since our study is concerned particularly with the interaction patterns that are relevant to cognitive development, the verbs we have identified are in that general area, and furthermore are appropriate for observing the interactions of adults and preschool children in a play setting. The lexicon would be somewhat different for other types of studies.

2. The category system for the vocabulary.

Information: One whole set of terms describes information exchange, giving, asking for and responding to requests for information

Statements	Requests for information	Responses to requests for information	no information utterances
labels	asks	answers	echoes
names (Fido)			
describes	asks if	replies	repeats
explains		acknowledges	
defines		reflects (questions back)	
contrasts			
compares		reminds	
indicates		hints	
clarifies			
corrects			
states			

discusses is used to indicate verbal exchange that for some reason cannot be described more precisely.

Explanatory behavior: A second major area of cognitive behavior is problem solving. There are many words that deal with different aspects of problem solving. In non verbal exploratory behavior for example, there are terms that describe aimlessness, and others that describe more or

less attentive goal directed behavior (see category system for nonverbal exploratory behavior), and others that describe comfortable enjoyment of well learned acts.

Aimless	Goal directed	Comfortable
wanders	approaches	plays
fiddles		uses
glances		
handles	manipulates	
	assembles	
	constructs	
	experiments	
	looks	
	examines	
	searches	

Attention: Another feature of many of these words is that they describe attentive behavior. Examines, searches, and experiments emphasize this feature. Other words explicitly describe attention, lack of it and requests for it.

Attention	lack of attention	requests for attention
points to	glances	alerts
searches	shifts attention	attracts attention
examines	fiddles	directs attention
holds attention		commands attention
attends		distracts

Help and encouragement: Problem solving involves both the ability of the individual and his effort. The same distinction is found in helping: sometimes help is directed toward solving the difficulty or increasing the child's skill, and at other times it is directed toward increasing his effort through encouragement.

Help	Encouragement
helps	coaxes
instructs	encourages
demonstrates	urges
approximates	challenges
guides (physical guidance)	describes task as easy
joins, works together with	describes task as hard
performs for	discourages
refuses help	
shows	

Ability: Sometimes the difficulty with the task is attributed to the task itself; sometimes it is attributed to the person. Thus the child may describe his inability or his skill, expressed by such terms as says he can, or says he can't, or the narrator may express the fact by states inability, or doubts his ability, or expresses confidence. In other circumstances the child or mother may state difficulty of the task. When the mother attributes difficulty to a task, she may be encouraging effort, or may be enhancing the child's feeling of success. These can be distinguished from the context.

Effort: Motivation and effort are expressed by such phrases as tries to or wants to, perhaps modified by adverbs like very hard or again when appropriate. Verbs like searches are also marked for effort in the vocabulary.

Success and failure: Finally, the outcome of effort is indicated by verbs like succeeds or fails,or can be expressed by using adverbs like successfully or correctly embodied in the description of the behavior itself. Success can also be indicated by specific verbs like fits for a puzzle piece. All of these various ways of noting the success of a goal-directed action when the overcoming of some difficulty is implied are given a success marker, either in the lexicon itself or the marker is attached to the appropriate word when success is implied in the context. The computer program will do the latter.

While the research is primarily concerned with information gathering and problem solving, it is clear that other aspects of interpersonal interaction cannot be ignored. There are two other areas, first the expression of feeling and the provision of emotional support, and second the area of behavior control: commands, permissions and the responses to them.

Expressions of feeling or evaluation

expresses (any emotion)
positive or negative

Positive	Negative
approves	criticizes
comforts	degrades or derrogates
encourages	doubts

Behavior control: Behavior control is described in several classes of word

Requests for behavior

Positive	Negative
invites	questions
suggests	forbids
demands	prohibits
persuades	stops

coaxes	coaxes not to
insists	
forces	restrains
bribes	threatens

Sanctions

allows	restricts
permits	limits

Responses to requests for behavior

accepts	resists	
agrees (verbally)	refuses	disregards
complies (behaviorially)	dissents	

Responses to requests for permission

permits, allows, denies permission

Responses to compliance or resistence

ignores
praises
reproves
insists.

In addition to these verbs which describe interpersonal behavior and the particular acts involved in information gathering and problem solving, the observer uses many ordinary verbs to describe behavior itself: stands, sits, picks up, replaces, walks, goes, etc. These generally are easy for the observer to use and their meanings are clear. The program may need to label the verb as "picks up" even when the particle is separated from the verb (as in "The child picks the red block up."), but the observer is free to use the words that come naturally.

Finally, the vocabulary contains a list of nouns, many of which are specific to the play room. We provide the observer with a set of standard names for objects that might be confused or difficult to describe. There are, for example, three different kinds of latches on the lock box. One is the hasp lock, one the spring lock and the peg lock. Some human shapes that go on the magnet board are called magnet-man and magnet-woman to distinguish them from the doll-man and the doll-woman.

This vocabulary has been presented largely from the point of view of the observer who must describe the interactions that take place. This is his repertoire of nouns and verbs that supplements or refines ordinary language. The words have not been organized according to the features that are used in the analysis.

To illustrate how these words are dealt with in the analysis, we present Table 5. The features are listed on the left and the words across the top. A plus sign in any cell means that that word is marked for the feature on the left. The minus sign indicates not that the word is negative on that feature but merely not marked. These features are in the dictionary that is filed in the computer. The table lists only the inherent features of the words. In the course of the analysis, certain additional features are assigned on the basis of contextual rules. For example ask is assigned either an information request feature depending on whether the sentence says the mother asked the child to do something or asked him a question. Similarly features like response or success are assigned in the course of decoding the sentence if they are not inherent in the word itself. The process by which this assignment is made will be described later. At the moment we are interested in showing how the words used by the observer can be multiply classified into various semantic categories which are not apparent from thei listing.

Thus if we are looking in the analysis for an index of how supportive the mother is of the child, we can count all the acts of the mother which are marked supportive in the feature list. The coerciveness of requests can be indexed by the percentage of requests that are marked for coerciveness. The observer need not worry about all of these features, however. He can concentrate upon trying to produce as accurate and complete a description of the interaction as he can, and may use ordinary English to do so.

3. Scheme for Computer Analysis of Interactional Language Transcript*

A more-or-less typical portion of an interactional language transcript is shown in Table 6. In the most general terms, the computer analysis of a transcript involves reducing each sentence to a "canonical form". The reduction is achieved by eliminating non-essential words, assigning the remaining words to one of a set of analytical categories, and assigning semantic features to these words, based on the context in which they occur. "Context" in this sense means both the rest of the current sentence, and the neighboring sentences.

The grammatical analysis of a natural language sentence, in order to determine the deep structure, also involves a reduction to canonical form. This is done, for example, in IBM's Automated Recognition Grammar for English developed by Culicover et. al. (1969). Aside from the fact that the final canonical form is quite different, there are major differences in strategy between automated content analyses and a general

---

*
Paul Ward with Sharon Horner (dictionary compilation)

Table 5

Feature	Goes	Picks up	Asks	Demands	Suggests	Coaxes	Corrects	Confirms
Interpersonal	-	-	+	+	+	+	+	+
Verbal	-	-	+	+	+	+	+	+
Non verbal behavior	+	+	-	-	-	-	-	-
manipulative	-	+	-	-	-	-	-	-
locomotor	+	-	-	-	-	-	-	-
Request	-	-	+	+	+	+	-	-
Coercive	-	-	-	+	-	+	-	-
Information	-	-	-	-	-	-	+	+
Question	-	-	-	-	-	-	-	-
Response	-	-	-	-	-	-	+	+
Supportive	-	-	-	-	-	+	-	+

Features of a sample of verbs

Table 6

A portion of an Interactional Language Record

1. Mark and his mother are in the playroom.
2. Mark picks up a magnet letter.
3. His mother directs him to leave it.
4. They have emptied the puzzle.
5. The mother is trying to put it together.
6. Mark is lying on the floor.
7. He removes the bed from the house.
8. The mother directs him to put it back into the house.
9. The mother asks Mark what the farmer is doing.
10. Mark is lying on the floor.
11. The mother commands him to get up.
12. The mother picks up the tail.
13. Requests a label.

grammatical analysis. There are certain built-in constraints and regularities in the situation which one can capitalize on to make the content analysis much simpler.

The observer of interpersonal interactions is verbally reporting a series of external events, and this largely constrains the type of sentence which he will use. Declarative sentences, in which the first noun or nouns are the grammatical and logical subject, and the first verb is the main verb of the sentence, are the rule. The positions and forms in which embedded sentences occur are also quite limited. The construction "the mother asks the child to (verb) --" is easy and natural and covers a wide class of potentially observable behavioral situations.

The system of rules by which the analysis is carried out is quite flexible and can be changed from analysis to analysis. This s suggests that the rules can be written to cover not only the built-in grammatical regularities of the situation, but also the "idiosyncratic" individual grammatical regularities of the observers. A construction used by an observer, even if complex and "ungrammatical" in the strict sense, is potentially analyzable if used regularly and consistently. For example, sentence 13 in Table 6 has "the mother" as an implied subject, and repeated use of this device could be covered by the rule "a sentence which begins with a verb has an implied subject which is the same as the subject of the previous sentence."

In addition to intra-sentence regularities, there are some empirical inter-sentence regularities which can be tapped to aid in the analysis. These may or may not be of general linguistic relevance, but are quite valid within the restricted universe of the narrative behavior record. For example, in Table 6 the pairs of sentences 2 and 3, 4 and 5, and 7 and 8 are subject to the "it rule" which is simply "the referent of it is the noun which is the direct object of the previous sentence." A slight elaboration of this rule makes it even more general. In constructions such as "the mother asks the child -- the mother repeats it," the "it" refers not to a noun but to the previous sentence as a whole. However, the class of verbs which use "it" in this way are not the same class which use "it" as a noun referent and can be "marked" accordingly.

Finally, a content analysis can make certain assumptions about analyzability which a general grammatical analysis cannot. There is no need in a content analysis to determine whether a given sentence is grammatical or not; the sentences are assumed to be grammatical. In addition, since an observer will produce many hundreds of sentences in the course of a half-hour observation, it is not necessary to achieve one hundred percent analyzability of the transcript. An occasional "irregular" sentence to which the rules cannot apply is not important, since the content analysis involves counting occurrences of various categories, and a certain amount of variance in the counts is assumed in the analysis.

The canonical form for the analyzed output consists of five major categories. (The mnemonics following the categories are the notation used in the computer output). The categories are: actor (ACTR), act (ACT), recipient (RECP), content (CONT), and modifiers (LOCT or some other symbol.) The actor of an interactional language sentence is a noun naming one of the participants in the interaction. The act is a verb describing an action (or state) of the participant. The recipient, if present, is also a noun naming a participant to whom the act is directed. The modifiers, if present, are nouns used as objects of prepositions (e.g. on the table) or adverbs (e.g., angrily). The content can be a noun used as direct object, as "truck" is the content of "the mother hands the child the truck," or a content clause (CTCL) used, for example, as a noun or verb complement. "Child moves the truck" is the content clause of "the mother asks the child to move the truck." In this case the content clause had to be partially reconstructed, since its subject had been deleted as part of the grammatical transformation which produced the original imbedded sentence. The reconstruction just cited involved inserting the recipient of the main sentence as the actor of the content clause. In other cases, the actor of the main sentence becomes the actor of the content clause, as in "the mother tries to move the truck." Sentences of this type, however, can be distinguished because they do not have a recipient, and because the main verbs are drawn from a characteristically different set. Content clauses may represent a past act of one of the participants, a hypothetical future act as in the last example, or a fact about the physical world. Content clauses have their own actor, act, etc., although in "fact" clauses, the actor has a somewhat different meaning (e.g. "fire engines are red".)

Simultaneously with assigning the analytical categories to certain words in the sentence, the contextual semantic features are assigned. These contextual features are predicated of a word just in case it appears in a certain context in a particular sentence. They are to be distinguished from intrinsic features, which are predicated of a word in any context, and are part of the dictionary entry for the word which is input to the computer. The word "get" is intrinsically a verb. In sentences where it is used as the main verb, without a particle, it has the semantic feature "acquisition". When it is used with the particle "up", however, it acquires the semantic feature "locomotion." This could be indicated in the analysis by deleting "up" and marking "get" with the feature "locomotion," or by replacing "get up" by "arise" which has "locomotion" as an intrinsic feature. Similarly, "asks" is intrinsically a verb and has the semantic feature "request." When it is used in the context "-- asks where the truck is," it can be marked "information request", and when it is used in the context "-- asks the child to stand up" it can be marked "behavior request." It is of course arbitrary to say that "asks" is an information request in a certain context; one could as well say that the entire sentence is an information request. In either case

there is implicit semantic information in the structure of the sentence which must be made explicit in some way. The feature marking system is simply one convention for doing this.

The intrinsic features are indicated by means of a dictionary which forms part of the input to the analysis program. A program has been written in the Snobol4 computer language in order to compile a dictionary of the words in a representative sample of sentences from a number of transcripts. Snobol4 is a language used mainly for string manipulation. If the string is a sentence, it can be broken down into words. The language has a list processing feature. It is assumed that the dictionary of words approximates a dictionary which would be obtained were all the sentences of each transcript analyzed. The frequency of occurrence of the various words would be proportional. The program performs several functions. Each discrete word is extracted and a count obtained of its occurrence in the sample. The words are arranged in alphabetical order. If desired, it is possible to delete from the dictionary those words of low frequency which are considered to be unimportant.

Feature addition may also require looking at previous sentences for pronoun referents or expansions of imbedded sentences. For example, "the mother ignores the child's question" is handled by looking for a "the child asks--" sentence and copying it in as the content clause to replace "question." Some of the many verbal cues which an observer uses to tie together sentences in a narrative may be examined in terms of transformational history. The analysis of a transformed sentence like "I bought John's car" involves postulating a <u>hypothetical</u> pair of sentences "John has a car" and "I bought the car" which are the historical determinants of the sentence. In a coherent narrative record, however, a sentence often has a <u>real</u> transformational history. "Mother ignores child's question" implies the occurence of a sentence like "the child asks --" somewhere previously in the transcript. A complete analysis of the transcript demands a detailed understanding of verbal cues of this sort, since many relevant semantic features refer to behavior sequences which occur over several sentences of the observer. We would like to know, for example, when the act described in one sentence is a response to an act described in a previous sentence. This is in general a complex function of both the individual words chosen and the inter-sentence context.

The full set of instructions for analyzing a transcript is written in a rule shorthand which is fed to the computer, and which is described in detail in Appendix B. The operation of the rule system relies on the ordering of the rules as well as on the content of the individual rules, since each rule operates on the sentence as transformed by the previous rule. Table 7 lists in proper order the English equivalents of the rules which were used to analyze the transcript in Table 6. It is assumed that the words "the" and "a" were

deleted prior to the analysis. A few of the rules are "ad hoc" and of limited generality, but most of them are quite general and could be used to analyze satisfactorily large portions of a transcript. The actual rule system to analyze an entire transcript will have to take into account many grammatical forms not encountered in this small sample, and will consist of a much larger number of rules.

Table 8 lists the features used in the analysis together with their mnemonic symbols, and Table 9 reproduces the actual computer output for the analysis of the transcript.

Table 7

Rules for the Analysis of an Interactional Language Transcript

1. "Is" followed by the gerund form of a verb; rewrite with "is" deleted and the verb marked progressive.

2. "Have" followed with the past participle form of a verb; rewrite with "have" deleted and the verb marked perfect.

3. "Are" followed by a word which is not a verb; rewrite with the marking "modal" removed from "are."

4. Rewrite "picks up" as "lift", "get up" as "arise", "put back" as "replace", and "put together" as "assemble."

5. Rewrite "he" and "him" as "Mark" and delete "his".

6. A sentence beginning with a noun marked "person" followed by a verb; mark the noun "actor" and the verb "act".

7. A sentence beginning with two nouns marked "person" separated by "and" and followed by a verb; delete "and", mark the two nouns "actor" and the verb "act."

8. A string of two nouns marked "physical object"; delete the first noun.

9. A sentence beginning with "they" followed by a verb; search the preceding sentences for two nouns in sequence marked "actor" and substitute them for "they"; mark the verb "act".

10. A sentence beginning with a verb; insert the noun marked "actor" from the previous sentence and mark the verb "act."

11. The word "to" followed by a verb; delete "to" and mark the verb "infinitive."

12. A word marked "preposition" and "locative" followed by a noun; delete the preposition and mark the noun "locative".

13. A verb which is not marked "modal" followed by a word marked both "verb" and "noun"; delete the "verb" marking from the second word.

14. "It" in the current sentence; replace with the direct object of the previous sentence.

15. A verb marked "request" followed by a verb marked "infinitive" and "behavior"; mark the first verb "behavior request."

Table 7 (continued)

16. A verb marked "request" followed by a relative pronoun; mark the verb "information request."

17. A noun following a verb, and marked "person" and not marked "locative"; mark it "recipient."

18. A verb not marked "behavior" followed by a noun marked "recipient" followed by an infinitive; insert the "recipient" noun, marked "actor," in front of the infinitive, mark the infinitive "act", and mark all of the rest of the sentence from the new "actor" with "content clause".

19. A noun marked "actor" followed by a verb marked "behavior" followed by an infinitive; insert the "actor" noun in front of the infinitive and proceed as in 18.

20. "What" followed by a verb followed by a noun; delete "what", insert "something" after the verb, mark the noun "actor" the verb "act", something "content", and the whole thing "content clause."

21. Mark any remaining nouns following verbs and marked "physical object" and not "locative" with "content."

Table 8

Elements of the feature system

Grammatical	
pronoun	PRON
preposition	PREP
relative pronoun	RLPR
verb	VERB
noun	NOUN
present singular	PRSG
present plural	PRPL
past participle	PPRT
gerund	GERD
modal	MODL
progressive	PROG
perfect	PERF
Semantic	
behavior	BEHR
request	REQS
behavior request	BHRQ
information request	IFRQ
stasis	STAS
goal directed activity	GDAT
manipulation	MANP
locomotion	LOCM
assembly	ASSM
physical object	PHOB
environmental object	ENVR
toy in playroom	TOY
Control	
end of sentence	EOST
end of transcript	EOTR
derivative form	DVFM

Table 9

Computer Analysis of Interactional Language Transcript

Mark	ACTR	NOUN	PRSN				
Mother	ACTR	NOUN	PRSN				
Are	STAS	ACT	PRPL	VERB			
Playroom	EOST	ENVR	LOCT	NOUN	PHOB		
Mark	ACTR	NOUN	PRSN				
Lift	MANP	ACT	PRPL	VERB	BEHR		
Letter	EOST	CONT	TOY	NOUN	PHOB		
Mother	ACTR	NOUN	PRSN				
Direct	DVFM	ACT	PRSG	VERB	BHRQ	REQS	
Mark	RECP	NOUN	PRSN				
Mark	CTCL	RECP	ACTR	NOUN	PRSN		
Leave	LOCM	CTCL	ACT	PRPL	INFT	VERB	BEHR
Letter	EOST	CTCL	CONT	TOY	NOUN	PHOB	
Mother	ACTR	NOUN	PRSN				
Mark	ACTR	NOUN	PRSN				
Empty	DVFM	MANP	ACT	VERB	BEHR	PERF	
Puzzle	EOST	CONT	TOY	NOUN	PHOB		
Mother	ACTR	NOUN	PRSN				
Try	DVFM	GDAT	ACT	VERB	BEHR	PROG	
Mother	CTCL	ACTR	NOUN	PRSN			
Assemble	MANP	CTCL	ACT	PRPL	INFT	VERB	BEHR
Puzzle	EOST	CTCL	CONT	TOY	NOUN	PHOB	
Mark	ACTR	NOUN	PRSN				
Lie	DVFM	STAS	ACT	VERB	BEHR	PROG	
Floor	EOST	ENVR	LOCT	NOUN	PHOB		
Mark	ACTR	NOUN	PRSN				
Remove	DVFM	MANP	ACT	PRSG	VERB	BEHR	
Bed	CONT	TOY	NOUN	PHOB			
House	EOST	TOY	LOCT	NOUN	PHOB		
Mother	ACTR	NOUN	PRSN				
Direct	DVFM	ACT	PRSG	VERB	BHRQ	REQS	
Mark	RECP	NOUN	PRSN				
Mark	CTCL	RECP	ACTR	NOUN	PRSN		
Replace	MANP	CTCL	ACT	PRPL	INFT	VERB	BEHR
Bed	CTCL	CONT	TOY	NOUN	PHOB		
House	EOST	CTCL	TOY	LOCT	NOUN	PHOB	

Table 9 (continued)

Mother	ACTR	NOUN	PRSN					
Ask	DVFM	ACT	PRSG	VERB	REQS	INFRQ		
Mark	RECP	NOUN	PRSN					
Farmer	CTCL	ACTR	TOY	NOUN	PHOB			
Do	DVFM	CTCL	ACT	VERB	BEHR	PROG		
Something	EOST	CTCL	CONT					
Mark	ACTR	NOUN	PRSN					
Lie	DVFM	STAS	ACT	VERB	BEHR	PROG		
Floor	ENVR	LOCT	NOUN	PHOB				
Playroom	EOST	ENVR	LOCT	NOUN	PHOB			
Mother	ACTR	NOUN	PRSN					
Command	DVFM	ACT	PRSG	VERB	BHRQ	REQS		
Mark	RECP	NOUN	PRSN					
Mark	CTCL	RECP	ACTR	NOUN	PRSN			
Arise	EOST	LOCM	CTCL	ACT	PRPL	INFT	VERB	BEHR
Mother	ACTR	NOUN	PRSN					
Lift	MANP	ACT	PRPL	VERB	BEHR			
Tail	EOST	CONT	TOY	NOUN	PHOB			
Mother	ACTR	NOUN	PRSN					
Request	DVFM	ACT	PRSG	VERB	REQS			
Label	EOST	EOTR	PRPL	NOUN				

MOTHER INTERVIEW AND CHILD MEASURES

Though the primary focus of the research has been to study cognitive aspects of naturalistic mother-child interactions, we did want some of the mother's perceptions of her child's cognitive abilities, and some independent measures of the child's language development. We also wanted the mothers to have a chance to talk to us about the research. Much of these data are not analyzed at this time.

The interview was designed to tap several sorts of information.

1. Family composition.

2. The mother's evaluation of the naturalness of the child's activity and language in the playroom.

3. The mother's perceptions of the child's language development.

4. The child's usage of various grammatical forms.

5. The child's memory for and anticipation of events.

6. Language interaction in the family and with peers.

7. Play activities.

Finally, the mother had an opportunity to ask questions about the research. The questions were presented in approximately the same order in each interview. If the mother spontaneously discussed items which we planned to ask about later, we obviously omitted these questions later. The interview was semi-structured, leaving the interviewer free to phrase the questions in what seemed a comfortable way and also attempting to create an atmosphere of a discussion about the language development of the child rather than a formal question-answer period. Interviewers varied in their ability to create this atmosphere.

The interview was modified for the older longitudinal children in accordance with possible age changes.

A copy of the first mother interview illustrates these points. (See Appendix F).

Child Measurements

While the central purpose of the study was to study language in naturalistic situations, it also seemed valuable to assess the child's language in a more standardized situation. Not only is the test situation psychologically different from the naturalistic play room interaction, but in a test situation the limits of the child's vocabulary and grammar can be explored. In the play room he may speak more simply than he is capable of.

The selection of appropriate measures was not simple. Different measures were experimented with, until finally in the later sessions three language measures were administered. All measures were administered individually (usually with the mother present until age 4).

1. In order to assess the child's ability to recognize and label familiar objects, the picture vocabulary test from the Stanford-Binet was administered with the standardized instructions. The following objects are pictured (one to a page) and the child is asked to label them:

airplane
telephone
hat
ball
tree
key
horse
knife
coat
ship
umbrella
foot
flag
cane
arm
pocket knife
pitcher
leaf

2. In order to study the child's ability to use various grammatical transformations, a sentence imitation task developed by Frank and Osser was administered. The child was asked to repeat a series of sentences that had been selected to involve several different transformations. All the sentences were seven words in length and used a simple vocabulary.

The imitation task was designed to test the child's control over specific syntactic structures. The score therefore is concerned with the accuracy of the imitation of these structures only. Other errors were ignored. In each of the seven sentences certain words were designated as critical for correct imitation. The words so designated composed the Critical Structure. It was only accurate repetition of this Critical Structure of each sentence that determined the score. Critical Structures are underlined on the test blank shown in Appendix E.

3. The Wugs test was developed by Jean Berko (1958) to test the child's understanding of morphological rules. It was deliberately designed (with two exceptions) to use nonsense words to that the child could not answer by merely repeating a familiar word. Each item is composed of two sentences, one which is incomplete. Each is accompanied by a picture illustrating the noun or verb being asked about. For example, the first item on the measure is:

Picture of a Wug

"This is a Wug."

Picture of two Wugs.

"Now there is another one."
"There are two of them."
"There are two _____."

Items for this measure may be found in Appendix F.

These three measures take about 15 minutes to administer. They were recorded by the tester and they were also observed and recorded by an observer behind a one-way screen. Scores are based on the pooled assessment of the child's responses by the two scorers.

The Negro Dialect

Since 1965, we have come to learn much about the systematic nature of the grammar of the American Negro dialect. (Stewart, 1964, Loban, 1966, Labov, 1968) This work is particularly relevant to the scoring of all the measures described above. Such a characteristic as the absence of the third person singular markers on the present tense of verbs makes it impossible to decide whether the child lacks a particular grammatical rule or whether he is simply using his own dialect. In general, we tried to take into account the relevant information about the non-standard dialect when scoring the responses of our Negro children. So, for example, "a han' for hand," "hisself" for himself, sweepin' for sweeping were scored correctly even though the articulation was not in standard English. The scoring becomes almost impossible on the Wugs measure where seven items require consonant clusters marking plural, past and possessive which are typically not present in the child's non-standard dialect.

In addition to these incompatibilities between the tests and the children's language, the investigators were very impressed by the striking difference in the general language productivity of the Harlem children in the test situation and in the play room. Though the experimenters were skilled in the administration of the measures, and often the child's mother was present in the room, the children from Harlem were very reticent and unresponsive in the test situation. The children of Washington Square were quite different in the test situation, whereas the children from the two samples were much alike in the play room. The Harlem children, particularly on the Wugs test, seemed quite unsure how to respond to a nonsense sentence completion task. They frequently refused to answer, or whispered so softly that the responses were inaudible even to the person sitting directly beside them. Particularly when the items became difficult, the Harlem children seemed to withdraw whereas the Washington Square children were more willing to give a try. After extracting almost no language from a child in fifteen minutes, it was very illuminating to hear him go into the play room and begin to chatter away.

Labov has also reported very similar findings with older Harlem children. It seems very likely that some of the retardation on intelligence tests is an artifact due to this style of coping with an anxiety inducing test situation administered by an adult. When we asked the mothers what they thought about such a performance, they would sometimes interpret it as "putting us on". And we thought we could detect a certain note of support for the child's attempt to do so.

White children can also show a reluctance to try difficult items, others also clam up, others also appear distractable, but in our experience with children we have been able most of the time to obtain what we felt were valid test scores with many kinds of children. The Harlem children were different, not less intelligent, but more frustrating to an examiner.

To sum this up, we are convinced that data based only upon structured measures of language ability may give a very biased view of the child's actual language ability and productivity.

Summary

In this chapter we have described the sample, and how it was obtained and the experimental situation. Of the various measures four are based on actual interaction:(1.) VINEX,(2.) Syntactic description; (3.) Coding of exploratory behavior; and (4.) Interactional Language. In addition the mother was interviewed after each session and the child was given three verbal tests: (1) Picture vocabulary; (2) Sentence imitation; and (3) "Wugs" test of grammatical inflection.

Chapter 3

Results

The results of these studies can be briefly stated. There are clear general patterns of mother-child interaction whether that interaction is viewed through the VINEX ratings, or syntactic complexity. While there are significant age differences and significant differences between different subsamples, all of them are relatively small deviations about a general pattern.

This section of the report will emphasize this overall pattern. Following sections will report sample differences and age differences.

Amount of Interaction

The best measure of the amount of interaction is the total number of utterances. There are about seven hundred utterances in the interactions of the mother and child over a thirty minute period, although the range is from 52 to 1185. The mothers in every session average more utterances than the child and the difference is significant in most cases.

As we will see there are age changes in this ratio, but generally speaking mothers make 55 to 60 percent of the utterances. There is also a positive correlation between the mother's and the child's utterances . In most instances this correlation is significant. Thus it appears that the mother-child pair has a level of verbal interaction that is characteristic of the pair. Mothers who talk a lot have children who talk a lot.

This correlation is probably a result of the general norm that in a one-to-one situation, utterances are generally responded to. Since the unit utterance in this study is not a complete speech, but a sentence, many utterances are not responded to individually, because some of the sentences are within a single speech. Even so 35 to 40% of the utterances are responded to. Thus the pattern is one of interaction back and forth in which both people participate about equally but this interaction operates at different levels for different mother-child pairs.

This finding may appear obvious, but an alternative pattern might be that members of the pair compete for the floor and that if one person makes many utterances the other person makes fewer. Another possible pattern in mother-child interactions is that the mother encourages the child to talk. If he is silent she stimulates a conversation, but if he talks a lot, the mother remains relatively silent. The pattern that we generally found was an interactive one with conversation flowing back and forth between the two participants.

Codability of utterances and amount of understanding between the participants.

Not every sound uttered by the mother and child during the sessions could be coded. Some of the failure was due to the actual sounds themselves. Utterances like "choo-choo" were uncodable. More important were sounds that could not be transcribed from the tape recording and could not therefore be coded. Of these some were due to bad electronics or traffic noises, but a fair amount was due to incomprehensibility of the subject's speech. The empirical findings support the latter interpretation.

In the first place mother's utterances are uniformly and significantly more codable than the child's utterances. Furthermore, the codability of the child's utterances increases with age. Generally speaking the mother's utterances are 90% codable and the child's by age 3-1/2 are 80% codable.

As every one knows, children's speech is more easily understood by the child's own mother than by a stranger. In addition the Harlem mother-child pairs frequently spoke in a Harlem dialect and in these cases we suspected that some of the uncodability was a result of the transcriber's inability to understand the dialect. During part of the study we had a Negro transcriber and at other times we had a Negro assistant who checked the transcription against the tape in order to check the accuracy of the transcription and to fill in the blanks as much as possible.

We wanted, however, to try to assess how well the members of the pair understood each other. One such index seems to be the frequency of utterances coded QQTC (questions requesting clarification about transient information). Questions like "huh" or "what did you say" would be coded in this fashion -- although they are not the only ones. Several bits of evidence support the use of QQTC as an index of understanding or smooth communication. For example 17% of the questions are coded QQTC while only 3% of the statements concern clarification of transient information (SSTC). Another bit of support comes from the fact that questions requesting clarification about permanent information (QQPC) account for only 2% of the questions whereas in general equal numbers of utterances involve permanent and transient information -- if anything the permanent information is higher than the transient. But among clarification questions the ratio is eight to one in favor of transient We believe therefore that the high frequency of such questions is at least partly due to one participant not understanding the other.

They are not all due to the mother failing to understand the child. In fact children consistently ask more such questions than the mother; the differences are consistent in every comparison but never statistically significant. As we will see there are significant age and sample differences in the frequencies of QQTC utterances.

Frequency of questions and statements

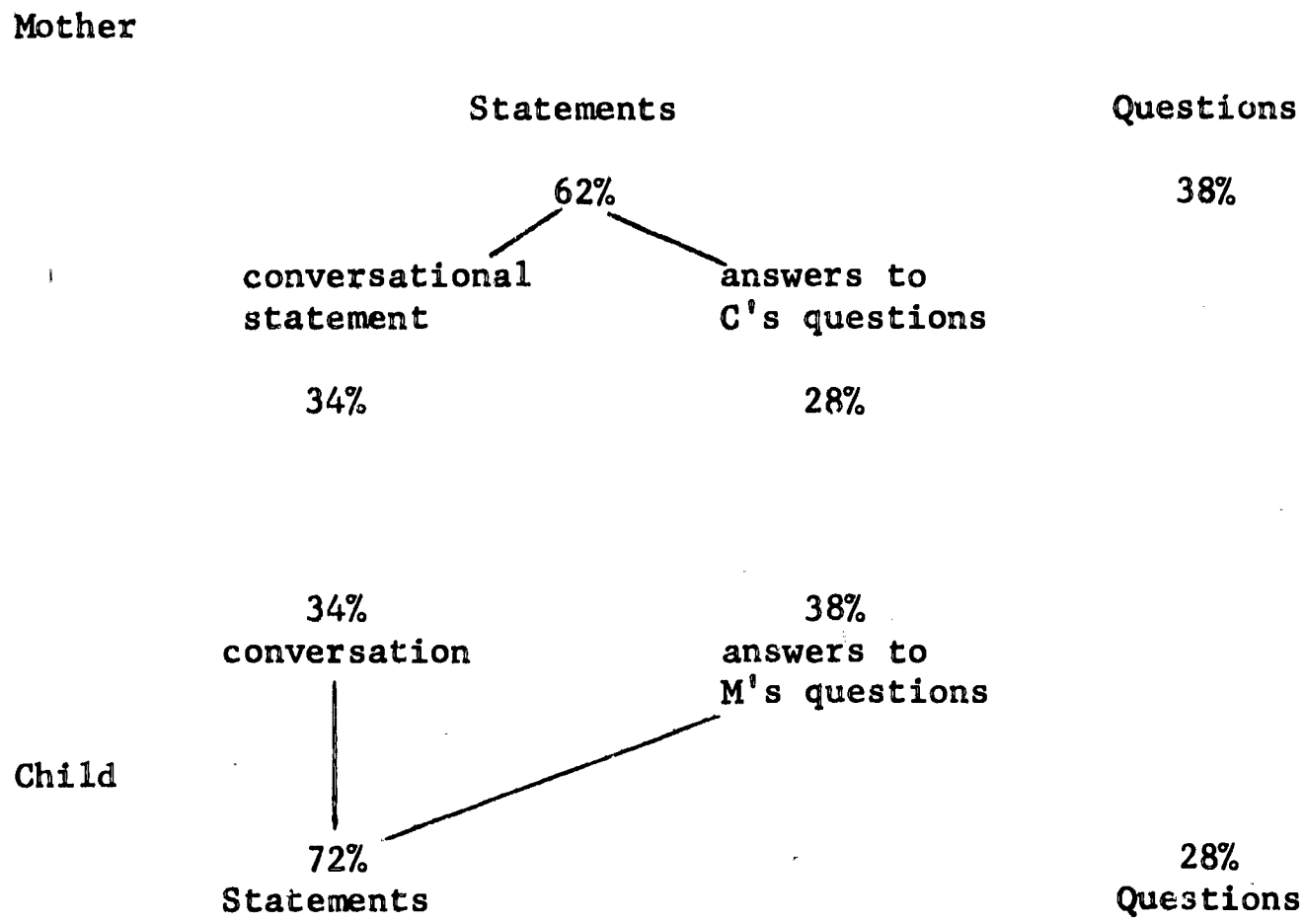
Generally speaking two thirds of the utterances are statements and one third are questions, and the percentage of the mother's utterances that are questions is significantly higher than the child's (see section on sample differences for further analysis). In many of the comparisons this difference is statistically significant, but the differences are relatively small. In no session is the average percentage of statements less than 55%.

It is interesting that the correlation between the mother and the child on these two variables is consistently negative, sometimes significantly so.

These findings suggest that the mother-child interaction takes on two forms, one a question-answer form, and the other an exchange of statements. Since the response to a question is much more likely to be a statement than a question, then an interaction that consistently took the form of questions by one participant and answers by the other would show a high percentage of questions by one participant and a high percentage of statements by the other. Inter-pair differences in the prevalence of such question-answer interaction would produce negative correlations between the frequency of statements or questions in the utterances of the mother and child. If the question asking was shared equally between the two participants, or if the interaction took the form of exchange of statements, then the differences in the frequencies of questions and the negative correlation would be reduced.

Thus a prevalence of statements overall may reflect the conversational type of interaction, the frequency of questions the question-answer type of interaction, and a mother-child difference in questions combined with a negative correlation might indicate the extent to which the question-answer interaction was initiated by one of the participants. These hypotheses can be firmly tested only by a sequential analysis of the utterances and we have not yet been able to make such a test, but the hypotheses seem plausible.

If we accept such an analysis then the average figures for the entire set of interactions are as follows:



Of the 62% of the mother's utterances that are statements, 28% are answers to the child's questions, leaving 34% that are in the conversational mode. This index, mother's statements minus child's questions, will be called the conversational quotient. Of the child's statements 34% are conversational and 38% are answers to the mother's questions. The agreement of the 34% conversational quotient for mother and child is an arithmetical necessity since statements plus questions total 100%. Thus 34% of the interaction is conversational while 66% is question and answer. Of the question-answer portion, the mother asked 10% more questions than the child.

This analysis is only roughly approximate. It assumes that the utterances of the mother and child are equal in total, and that every question by either party is answered by a statement from the other. Neither of these assumptions can be strictly accurate. but nevertheless an index composed of the percentage of mother's statements minus the percentage of the child's questions may give a rough measure of the amount of conversation, the sum of the mother's and the child's questions a measure of the question-answer mode of interaction, and the difference between the mother's and child's questions a measure of the onesidedness of the questioning.

Behavior Requests and Commands

Another indication of the nature of the mother-child interaction is the extent that it contains attempts by one participant to control

the behavior of the other through requests, commands and the like. There are several of the VINEX categories that bear on this aspect of the interaction. The category system for describing the form of the information (see P- 16) contains a category for behavior requests that includes questions and statements that ask the other person to do something or tell him to do something. These utterances almost always contain behavioral information in the content category and a certain number of them are categorized as commands in the category system for mode.

Between 10 and 30% of the utterances contain behavioral information but these are not all requests since they can include also statements about what the person is doing. Of these 2% to 18% of the utterances are coded behavior requests, and of these a very small number, 0% to 1.5%, are sufficiently coercive to be coded as commands. The code for behavior requests is probably the clearest indication of the amount of control exerted or intended to be exerted.

There is a highly significant difference between the number of behavior requests made by the mother and the child. In every comparison the difference is highly significant. Mother's behavior requests constitute 9% to 18% of her coded utterances, while the child's utterances contain only 2% to 9% of behavior requests. Thus it appears clearly that the mother is trying to control the behavior of the child by this means much more than the child requests behavior of the mother.

Content of the utterance

The information in the utterance was categorized as permanent information, transient information, behavioral information, fantasy information, and approval or disapproval. Of these the amount of explicit approval or disapproval was so small that no analysis was possible. Two of the other categories are particularly interesting.

Permanent information is information that has some permanent validity and therefore contributes to the individual's store of knowledge that might be useful in future situations. While there is no way of knowing that any particular bit of permanent information is new information, we presume that if a mother-child pair talked frequently about content that contained permanent information the child would have the opportunity to acquire valuable knowledge that would not be as available if the mother-child interaction contained primarily transient information.

The data indicate that in this kind of free play interaction about one-third of the utterances contain permanent information. The differences between the mother and child are never more than a few percentage points and generally the value for the child's utterances is greater than the mother's. In some instances this difference is significant. This difference seems to be a result of the fact that there is consistent and generally significant difference on behavioral information. Mother's utterances contain more behavioral information.

This is consistent with the previous finding that mothers make more behavioral requests than children.

Fantasy information is one variable on which there are age and sample differences (see P. 17). In general fantasy information is not frequent (less than 10%) and consistently there is more fantasy information in children's utterances than in the mother's. The differences are not always significant for individual sessions, but the difference is consistent in all eleven of the comparisons tested. There is also a high positive correlation between fantasy in the mother and the child's utterances. In other words some mother-child pairs engaged in fantasy play and this raised the amount of fantasy information in both of their utterances, but the child talked more about the fantasy and the mother's role was an accessory one.

Mode of information exchange

The utterances were also coded into categories describing different kinds of information exchange like explanation, labels, description, command, etc. (see P 18 for coding manual). As would be expected these occur with quite different frequencies. In rank order the general frequency of these categories are:

Description	30 to 50%	Mothers higher than children
Labels	17 to 28%	Children higher than mothers
Specification	7 to 22%	Children higher than mothers
Clarification	4 to 20%	No consistent difference
Explanation	1 to 6%	Mothers higher than children
Expression of feeling	1 to 4%	No consistert difference
Limits	.2 to 4%	Too infrequent to be compared reliably.
Demonstrations	.2 to 2%	Too infrequent to be compared
Commands	0 to 1.5%	Too infrequent to be compared

As we will see (P 81) modes on which mothers are higher than children generally increase with age and vice versa.

Quality of the responses

The analysis of the responses to utterances is complicated by several factors. The final coding system identifies each utterance as response-demanding or not, and each utterance is either coded as a response and described or it is coded as not a response and the reason for its not being. This is a refinement in the coding system that was introduced late in the study and therefore we cannot utilize it to analyze carefully the structure of the dialog between the mother and the child. Early in the coding, many utterances were judged to be ignored because of the absence of a response. Later the definition of an ignoral was narrowed considerably and therefore the data on the

percentage of utterances ignored contains an artifact due to the coding change. This change took place between the coding of West Harlem sessions two and three. It is interesting and important to assess how adequately each person in the pair responded to the expectations of the other person for a response, but the data are unfortunately inadequate for a careful analysis of that question. We can say that 30% to 40% of the utterances were responded to, and that only 1% to 2% of the utterances were clearly ignored (according to the latest coding). Whether all of the remainder were utterances that needed no response,or did not permit one, cannot be ascertained.

Questions are generally response demanding; about 70% of the childs questions were responded to. The child's responses to the mother's questions were consistently lower, especially at the younger ages,and that age trend will be discussed later. Statements are generally less response demanding and it is not suprising therefore that the percentage of statements responded to never averages above 40% for a session and is sometimes as low as 20%. Again there is a consistent tendency for the mother to respond more consistently to the child's statements than the child responds to the mother's statements. In general, however, one does not get the impression that either mother or child were unresponsive and as we will see this holds for all the samples.

How were the utterances responded to? The category system provides for three categories, direct response, peripheral response, and no-information response. This coding is more applicable to questions than statements. 25% to 50% of the questions were directly responded to; i.e. whatever was asked was answered. The next highest category is no information ranging from 10 to 30%; the least frequent type of response was a peripheral one which never goes higher than 10%. Interestingly, the mothers are more likely to give peripheral and no-information responses to questions than children; these are consistent differences across all sessions and frequently significant. The percentage of direct responses to questions shows no consistent pattern of mother-child differences; if anything the child gives more direct responses than the mother. Thus the mothers are generally more responsive than the child, but their excess of responses are peripheral and no-information responses rather than direct ones.

This should not be interpreted as being bad. Some of the no-information responses -- we cannot say just how many -- are attempts on the part of the mother to help the child find the answer for himself and represent an educational strategy, not just perverseness nor ignorance of the right answer.

The responses to statements show a different picture. There is a clear significant tendency for the mother to respond directly to statements more than the child does, and a less clear tendency for the same to be true of no-information responses. Peripheral responses to statements are very rare after age 2-1/2 and cannot be reliably compared.

Syntactic Complexity

Since one of the commonest hypotheses to account for the difficulty black lower class children have in school is the lack of adequate language models in the home, we felt it important to assess the level of syntactic complexity in the interaction.

We have available, therefore, a measure of the syntactic complexity of the mother's utterances in a conversation with an adult interviewer at the end of the first session, then measures of her complexity and the child's during the interaction itself for each of the sessions.

We experimented with five different measures of complexity, one, the number of words per sentence, second the complexity score derived from the Frank and Osser measure, third the average number of transformations per sentence, fourth, the total number of different transformations and fifth, the total number of different transformations divided by the number of sentences. Given that the individual samples contain the same number of utterances all of these are correlated highly, but as seen in Table 10 , for the five-year-old Harlem children,

Table 10

Complexity Correlations

N = 20 Group: Age 5 Harlem children

		2	3	4	5
1.	Mean Words	.87	.70	.68	.54
2.	Mean Complexity		.89	.59	.54
3.	Mean #T/Sentence			.45	.51
4.	Total Diff. Ts				-.07
5.	#Diff. Ts/Sentence				

r .05 = .369 r .01 = .503

both the total number of transformations and the number divided by the number of sentences are seriously distorted by the number of sentences in the sample. The correlation of -.07 is due to the presence in the sample of one child who talked very little. Therefore his total number of transformations is low, but his average number per sentence is high.

The number of words is obviously a good measure highly correlated with complexity and is easily calculated, but the complexity score is a more refined index. Therefore the data to be reported consists of complexity scores.

The following data shows the differences between the complexity of the mother's language in the interview, and the average complexity of her language in the interaction sessions, and the average complexity of the child's language in the interactions.

Mother in interview	13.33
Mother in interaction	4.89
Child in interaction	3.40

The general pattern is clear and consistent. The mother in speaking to an adult interviewer uses much more complex syntax than when she is speaking to her child in the play situation. Her grammatical complexity, however, is consistently and significantly greater than the child's in the same interaction. There are both sample differences and age trends in this measure which will be discussed later.

One caveat must be discussed in the interpretation of these data. The grammatical complexity of language in a naturalistic situation is not merely a function of the competence of the person talking but also of the situation itself.

We demonstrated this in several ways. For one thing the setting in the play room influences the complexity of the language used. Every utterance coded for complexity was also coded for topic, and for the toy being played with at that time. There is a significant effect of the object on the complexity for both mothers and children. Thus the mean complexity of a mother-child pair may reflect partly what the child played with during the session.

We also calculated the mean complexity of utterances in the various VINEX categories. Again there are large differences. In one session for example, questions requesting an explanation had an average complexity of 8.4 while at the other end statements coded labelling had an average complexity of only 3.4. Since some of the VINEX categories never appeared in the sample of one hundred sentences syntactically coded for an individual, it is difficult to make clear quantitative statements about the differences in complexity for different VINEX categories, but they are quite consistent. Behavior requests, for example, appear in the record of every subject in 12 of the 16 sessions on which we have data. Its rank in order of complexity ranges from 1 to 3 in different sessions. Questions are more complex than statements in every session coded. Permanent information utterances are always more complex than transient information utterances on the average over a session. There are five VINEX categories, questions, statements, permanent information, transient information and descriptive statements, that appear in the sample for every subject in every session. For these five the coefficient of concordance-roughly comparable to a correlation coefficient-across the 15 sessions is .79. Such a value is significant at some astronomical level, the value of F is 43.

Thus it is clear that the difference between the mother and child's level of complexity might be due to the different composition of their utterances. Fortunately we have clear evidence that the mother's complexity is greater than the child's on each of the five VINEX categories listed above. Both use more complex grammar in questions than any of the others, but the complexity of the questions is greater in the mother's than in the child's utterances.

Sample Differences

The data on sample differences are drawn from the analysis of the verbalizations during mother-child half-hour interactions in the playroom described in Chapter 2. The total sample consists of 62 children and their mothers. One subgroup of 12 came from the Washington Square area of New York, the remainder from a group of mothers and children from West Harlem. A sub-group of this Harlem sample, identified on Table 1 as "old longitudinal" were matched in sex and age with the Washington Square group. Comparisons of these two groups at 2, 2-1/2, and 3 years of age comprise one part of this section on sample differences. A description of the socio-economic and ethnic backgrounds of the families may be found in chapter 2. These groups will be referred to as WH (West Harlem) and WS (Washington Square).

In order to see if some of the differences <u>within</u> the West Harlem sample are reflected in differences in the mother-child interactions, the West Harlem sample was subdivided into two sub-groups reflecting certain differences in social class. As was shown in Chapter 2 the Hollingshead index of social class emphasizes father's occupation and the Harlem males are underemployed for their educational level. Therefore the division of the West Harlem sample was made entirely on the basis of the father's education, or when the father was absent on the basis of mother's educational level. The so-called upper group (HE) consists of families in which the father had completed high school and the lower group includes those families where the father had less than a high school education (LE).

Comparisons of these two groups of Harlem children based on educational level of the head of the family constitute the second set of sample differences in this sample.

The verbalizations were recorded during the play session, transcribed and coded according to the two measures described in Chapter 2, VINEX (verbal information exchange), and the Frank and Osser syntactic complexity measure (see Appendix A).

Probably the most striking finding regarding these verbal interactions is the relatively few significant differences between any of the sub-groups. This is particularly important in the light of the currently held opinions concerning hypothesized mother-child interactions in our poor families. In Chapter One, we have referred to a number of papers which suggest that the slum or ghetto or "disadvantaged" child has had inadequate experiences in verbal interactions with his family and peers. Often these are hypotheses, unsupported as yet by data.

In our analysis of sample differences, we find some differences which are puzzling to explain, but we have not found the pattern of low interaction which is often attributed to "ghetto" families either in contrast to the WS sample or in the contrast between higher and lower educational level (HE & LE) groups. There are many individual differences in the dyads of mothers and children in both groups but in this report we will deal only with differences in group means on a number of language variables.

Group Differences on Verbal Information Exchange(VINEX) categories

Table 12 on pages 67 and 68 lists the VINEX categories, and shows the variables on which there were significant differences between West Harlem (WH) and Washington Square (WS) mothers, and between the two groups of children. Significance was tested by means of t tests. Means, t's, and significance for each variable may be found on the Tables 1 through 37 in Appendix D.

Mean number of utterances

It is informative to look at the data on simply mean number of utterances over all sessions in the two groups.

Table 11

Mean number of utterances in
Washington Square and West Harlem samples

	Mothers			Children		
Session	WH	WS	Sig.of diff.	WH	WS	Sig of diff.
1	487	384	.05	397	306	NS
2	403	339	NS	345	320	NS
3	340	350	NS	297	328	NS

There were no significant differences between the two groups of children. Both children and mothers in the WH groups give more utterances than the WS groups in sessions 1 and 2. The only difference which reached significance was between the two groups of mothers in session 1 ($p < .05$), with WH mothers having more utterances than WS mothers. By session 3, WS mothers and children had a larger frequency of utterances but the differences were not significant.

Similarly there were no striking differences between the LE and the HE samples. There is a tendency for HE mothers to talk more than their LE counterparts, but this difference reaches significance only at age 5.

Table 12

Significance of Sample Differences[1] on VINEX Categories

	Mothers Session			Children Session		
Category	1	2	3	1	2	3
Mean Utterance Frequency	.05 (WH)	NS	NS	NS	NS	NS
Percent Codable Utterances	NS	NS	NS	.01 (WS)	.05 (WS)	NS
Form						
Questions	NS	.05 (WH)	NS	.01 (WH)	NS	NS
Statements	NS	.05 (WS)	NS	.01 (WS)	NS	NS
Information Content						
Permanent	NS	NS	NS	NS	.05 (WH)	NS
Transient	NS	NS	NS	NS	NS	NS
Behavioral	.05 (WH)	NS	NS			
Fantasy	NS	NS	NS	NS	.05 (WS)	NS
Mode of Information Content						
Explains	NS	.01 (WS)	.01 (WS)	NS	NS	NS
Limits	NS	NS	NS	NS	NS	NS
Clarifies	NS	.01 (WH)	NS	.05 (WH)	.01 (WH)	.05 (WH)

	Mothers Session			Children Session		
	1	2	3	1	2	3
Describes	.01 (WS)	NS	NS	NS	NS	NS
Feeling	NS	.01 (WH)	NS	NS	NS	NS
Demonstrates	NS	.05 (WS)	NS	NS	NS	NS
Commands	NS	NS	NS	NS	NS	NS
Labels	NS	NS	NS	NS	NS	NS
Specifies	NS	.01 (WS)	NS	NS	.05 (WS)	NS
Response Quality						
Direct	NS	NS	.01 (WS)	NS	NS	.05 (WS)
Peripheral	NS	NS	NS	NS	.05 (WS)	NS
No inf.	.05 (WH)	NS	NS	NS	NS	NS
M's utterances responded to	NS	NS	NS	NS	NS	NS

1. (WH) or (WS) below Significance level indicates group showing higher mean use of category.

Codability and Understanding

As indicated earlier, there is evidence that the codability of utterances is related to the comprehensibility of the children's speech. The WS children's utterances were significantly more codable than the WH children in sessions 1 and 2. To some degree this difference may be artifactual because the WS interactions were observed 6 months later than the WH interactions in real time, though the ages of the children were comparable. During that six months the electronics was improved and the staff became more practiced in transcribing the tapes. However, since the codability of the mothers' utterances is 90% in both the WH and the WS samples, the validity of the difference in the codability of the children's utterances is supported. There are no significant SES differences in codability.

We saw earlier that the prevalence of QQTC utterances (questions asking for clarification of transient information) may be an indicator of the lack of comprehension of one participant by the other. On this index there are large differences between the two samples. The percentage of QQTC in the WH sample runs between 18 and 35%, while for WS the range is 6 to 15%. Thus it appears that the WH mothers and children found each other's utterances less comprehensible than the WS pairs. If this difference in the comprehension of the mother's utterances is a valid one, it is not due to their actual comprehensibility, since the utterances of both groups of mothers were equally codable, hence comprehensible to observers and transcribers.

The comparison of the two educational groups on QQTC is interesting. The LE mothers are consistently higher in the amount of QQTC than the HE, but the direction of the difference is reversed in the utterances of the children. Both sets of differences are consistent, but in only one instance the difference is significant, see Tables 13, 14 and 15 .

Table 13

Percentage of utterances coded QQTC in the middle and lower educational level Harlem families

	Middle Class		Lower Class	
Age	Mother	Child	Mother	Child
3	19.41	31.02	22.45	22.36
4	13.47	30.11	19.02	15.91
5	9.63	15.64	12.23	14.54

Table 14

Educational level differences in
percentage of QQTC for mothers and children

(High educ. level -- Low educ. level)

Age	Mother	Signif.	Child	Signif
3	-3.04	t= .63 ns	+ 8.66	t= 1.36 ns
4	-5.55	t= 1.03 ns	+14.20	t= 2.51 5% level
5	-2.60	t= .7 ns	+ 1.10	t= .21 ns

Table 15

Mother-child differences in
QQTC for the two educational levels

Age	Mother-Child difference			
	HE	Signif	LE	Signif
3	-11.61	t= 2.13 ns	+ .09	t= .02 ns
4	-16.64	t= 2.15 ns	+3.11	t= .57 ns
5	- 6.01	t= 1.26 ns	-2.31	t= .70 ns

Another way to describe it is to say that with the lower educational level sample, the mother and child are about equal in QQTC (and this is also true of the WS sample), while in the higher educational level Harlem interaction, the child has many more QQTC's than the mother. See Table . If the mother were relatively didactic, rather than conversational, the child might be more likely to ask the mother to clarify than she would him. Perhaps therefore, HE Harlem mothers are more didactic than LE. This interpretation is supported by the finding that in the HE Harlem groups, mothers ask significantly more questions than children, while the differences in the LE group are much smaller.

Form of the utterance

Children

As is true at all ages on our sample, children in both groups and in all sessions tend to use statements more frequently than questions

in their verbal interactions with their mothers. There are clear sample differences, however, with WH children using questions more frequently than the WS group. These differences are significant at the .01 level in session 1, but by session 3, the groups are very similar in the form of their utterances.

Table 16

Mean percentage of children's use of statements vs questions

	Session	WH	WS	Sig. of diff.
Question	1	34	19	.01
	2	30	23	NS
	3	30	28	NS
Statements	1	66	81	.01
	2	70	76	NS
	3	71	72	NS

Mothers

Mothers' percentages of questions and statements are seen in Table . WS mothers used statements more frequently than the WH mothers in sessions 2 and 3, while questions occurred more frequently in the utterances of the WH mothers in sessions 2 and 3. These differences are significant at the .05 level in session 2 for both groups of mothers.

Table 17

Mean percentage of mother's use of statements vs questions

Questions

Session	WH	WS	Sig. of diff.
1	35	39	NS
2	45	35	.05
3	43	36	NS

Statements

Session	WH	WS	Sig. of diff.
1	65	61	NS
2	55	65	.05
3	57	64	NS

If one takes the theoretical assumptions concerning the conversation quotient proposed on page 5 (M's statements - Child's questions), the pattern of interaction is seen on Table to vary significantly between the two groups.

Table 18

Group differences in Conversational Quotient across 3 Sessions

	Percent of Mother's Statements	Percent of Child's Questions	Conversational Quotient
WH	59	31	28
WS	63	23	40

These differences are consistent across the three comparisons. Thus, the West Harlem pairs interacted more in terms of questions and answers than the WS pairs while the interaction between the WS mothers and children might be characterized as more conversational. There are no consistent differences between the two groups of differing education levels on this index of conversational interaction.

Behavior Requests

Behavior requests reflect the tendency of one person to ask the other to do something. Mothers make more such requests than the children in both the Harlem and Washington Square samples, but there are no consistent or significant differences between the two samples. For the educational level comparisons the most significant mother-child differences are in the LE group.

Information Content

One of the major interests in this research was to measure the amounts of various kinds of information which were exchanged in mother-child interactions. On the VINEX coding system (see p.) information content is categorized as permanent, transient, behavioral, fantasy, approval or disapproval. We have defined permanent information as that which adds to or reinforces a child's enduring fund of information about his world; transient as being of a more "here and now" quality; behavioral as purely information requests or statements concerning another person's behavior, and fantasy as the sort of information transmitted only through the child or the mother's verbalizations during play or pretending.

Findings:

A. The transcripts of the verbal interactions of both mothers and children in both [illegible]H and WS) contained approximately one-third permanent inform[illegible]nd [illegible]e-third transient information.

Table 19

Average use of permanent and transient information

Sessions 1, 2, 3 combined

	Permanent		Transient	
	WH	WS	WH	WS
Children	37	38	39	37
Mothers	36	39	35	33

Differences between the two samples in all three sessions are generally small and insignificant, though there is one isolated and confusing significant difference in session 2, when WH children's utterances contained significantly more permanent information than the WS group, ($P < .05$) and when LE children were higher than HE group.

B. Behavioral information, like behavior requests, occurred with about equal frequency in both groups in all three sessions. Mothers use this kind of information exchange more frequently than children, the range being 21 to 31% over all sessions for the children. The utterances of WH mothers contain significantly more behavioral information than WS mothers in session 1 only. By session 3, both groups are alike in the use of this category. The two groups of children do not differ in any of the three sessions.

C. Fantasy. The percentage of fantasy was low in both groups of children, with mean percentages ranging from 3.5 to 11.3% over the three sessions. Children in the Washington Square sample used fantasy more frequently than did the Harlem children in all sessions, the difference being significant at the .05 level in session 2.

D. Approval and disapproval Unfortunately, there were not enough utterances in which the coders felt the information contained was pure approval or disapproval, so we have no data on group differences on this category for either mothers or children. This is probably due in part to the fact that approval or disapproval was usually connected with a behavior request or statement, and since we did not double-code on the information categories, those qualitative aspects of the interaction were lost.

Mode of Information Exchange

Several interesting findings emerge from the analysis of the "mode" in which information exchange takes place between mothers and children in these play sessions.

A. Description and labelling occur with fairly high frequency at ages 2-1/2, 3, and 3-1/2 in the conversations of both the mothers and the children in both groups. The range of mean percentages for description is from 29-45% for the children's sessions; 35-53% for the mothers'. Percentages of utterances containing requests or statements involving labels range from 20-28% for the children, 17-22% for the mothers.

In sessions 1 and 2, the Washington Square children use description more frequently than the West Harlem children, but by session 3 there is little difference between the groups. The same pattern is true of the mother's use of description. Sample differences are significant at the .01 level in session 1.

There is no significant differences in the frequency of labelling in either the mothers' or children's groups.

B. The percentages of utterances either giving or requesting clarification are very different in the two groups. There is consistently more of this kind of verbal interaction in all three sessions of both the mothers and children from West Harlem. The differences are significant in all three children's sessions, and session 2 for the mothers. This difference is largely attributable to the prevalance of QQTC in the West Harlem records.

C. The other mode categories -- limits, feeling, demonstrates, commands, specifies, explanation -- occur with low frequency in all the sessions of both groups. There is a significant difference ($p < .01$, sessions 2 and 3) in the use of explanation, Washington Square mothers using this mode more frequently than West Harlem mothers. However, the range of percentage of this category over all 3 sessions is only 1-5% for children and 2-7% for mothers.

Response Quality

If the utterance was judged to be a response to a previous utterance, it was scored for the quality of that response, i.e. whether it was direct or peripheral, or a "no information" response, such as reflecting back a question to the person who asked it or saying "I don't know."

Findings:

A. Twelve to 20% of the mother's responses were direct responses, and in sessions 1 and 3, WS mothers gave more of these than WH mothers. The difference was significant at the .01 level in session 3. There were almost no differences between the groups in session 2. The range of children's direct responses was similar to the mothers, 15-22%, and again in session 3 WS children gave significantly more of them ($p < .05$). The same pattern of difference appears in both responses to questions and to statements. There were no SES differences within the Harlem group.

B. Peripheral responses, such as "ask your father", occurred with the least frequency, 2-6% in the two mother's groups, and 2-8% in the children's, and there are no significant differences between either the two groups of children or mothers or with SES.

C. No information responses occurred in 6 - 15% of the mothers' utterances, 11-18% of the children's. There were two differences which reached significance at the .05 level, WH mothers gave more of these responses in session 1, WH children gave more than WS in session 2.

D. We also calculated the percentage of mother's utterances and children's utterances which were responded to by the other person in the pair. 28-39% of child's utterances were responded to, and 26-34% of the mother's. At later ages we also scored an utterance as response-demanding, but unfortunately for these early ages we do not have that information, so we are unable to say what percentage of the utterances judged to require a response received one. There were no significant differences between either groups of mothers or children in the utterances responded to.

In general the, these data on sample differences confirm the previous statement that there are more similarities between the groups than differences. Where there are differences, they often tend to disappear by age 3-1/2. The variables where this is not true are: use of explanation, direct responses and especially clarifications. Another striking difference between the samples is in what we have called amount of conversation (M's statements vs children's questions) where the WS pairs show consistently more of this sort of interaction.

We are collecting data on these same children at age 5, and it will be interesting to see what pattern of verbal interactions emerges at that time. For one thing, we know from our cross-sectional 5 year olds, that the amount of labelling decreases and explanation increases with age, and it will be interesting to see whether the groups differ on these variables as well as the frequency of clarification as these children become older.

Syntactic Complexity

On syntactic complexity there are significant differences between West Harlem and Washington Square as shown in Table 20 .

Table 20

Sample Differences

Mean Complexity

	Mothers			Children		
Session	WH	WS	Diff.	WH	WS	Diff
1	4.3	5.1	.8*	2.7	3.1*	.4*
2	4.8	5.2	.4	3.0	3.6	.6
3	4.6	6.0	1.4**	3.2	4.1*	.9*

* difference significant at .05 level

** difference significant at .01 level

Consistently, the WH mothers and children speak in less complex grammar than do the WS groups. However, the actual differences are small especially in view of the complexity of the mothers speech in the interview. Table 21 shows how adult language reaches a mean complexity of 13 when mothers talk to an interviewer.

There is no difference in the complexity of the language of the two groups of mothers when talking with another adult, but there are highly significant differences ($p < .001$) in both groups between the mothers conversation with the interviewer and when they were talking to their children. Also in both groups, the mother's language is significantly more complex than the child's.

These results do not seem to describe the difference between the two samples as the difference between an elaborated and a restricted code. In general all mothers use a restricted code when talking to their children. This restriction makes them more understandable to the child and possibly teaches the child syntax. There are differences in the complexity of the speech of the two samples, but this is not due to the West Harlem mothers having less linguistic resources at their command. With the adult interviewer their language as measured by the Frank and Osser system is just as complex as the Washington Square sample.

Table 21

Average Complexity of Mother and Child

	Harlem		Washington Square
Mother (Interview) [1]	13.2	=	13.4
Mother (Interaction with Child) (average of 3 sessions)	4.6	<	5.4
Child (Interaction with Mother) (average of 3 sessions)	3.0	<	3.6

> sign at .05 level

>>> sign at .001 level

1. children were 2-1/2 years of age - Session 1.

If it is not due to linguistic resources, what is the cause of the difference between the two samples? It is partly due to the difference in the style of interaction. West Harlem mothers, for example, ask significantly more clarification questions than Washington Square mothers, and the average complexity of such questions is low, 3.2 compared with an average of 4.3 for that session. The Washington Square mothers are higher on fantasy and on explanation, and both of those categories involve more complex syntax regardless of which sample uses them. So an apparent difference in syntactic complexity could be an artifactual result of differences in other features of the mother-child interaction.

This is not the whole explanation, however, because many of the differences between Washington Square and West Harlem remain even if the VINEX category is held constant. Clarification questions are relatively simple sentences, but their average complexity for the WS sample in session 1 is 3.9 while for the WH sample it is 3.0. Similarly for many other categories, the sample differences hold up.

There seems to be, therefore, a real but small difference in the syntactic level at which the mother-child pairs in the two samples operate despite the equivalence of their language when talking to the interviewer. The difference is not large enough, however, to justify the conclusion that the WH children are deprived of meaningful language experiences or that the WH families speak in some very primitive way.

Age Differences

While significant sample differences were few and far between, there are a number of indications that mother-child interaction changes with age. VINEX differences as a function of age were analyzed for both West Harlem (WH) and Washington Square (WS). The WS age differences are based on the longitudinal study which included 1/2 hour samples of mother-child interaction at ages 2-1/2, 3 and 3-1/2. The WH differences are based on two separate analyses, one on the longitudinal sample of five ages (2-1/2, 3, 3-1/2, 4 and 4-1/2) and the other on a supplemental sample of ages 3, 4 and 5. The primary purpose for supplementing the WH sample was to increase the reliability and validity of the findings by increasing the number of subjects observed at certain ages. This was accomplished by adding 10 non-longitudinal mother-child pairs at ages 3 and 4 and at the same time add a new sample of 20 mother-child pairs at age 5.

The VINEX categories analysed for age differences were the same as those previously discussed. Age differences on each of these categories will be reported for both mothers and children within each of the three groups defined above. The means for each level for each group and the significant differences between ages for both mothers and children, are presented in Table 41 in Appendix D.

Frequency of Utterances

There is a general tendency for the mothers to talk less to their

children during the interaction as the child becomes older. This is true for both the WH and the WS samples. This decrease with age is sharpest in the lower class Harlem sample. The frequency of children's utterances is however more mixed. There are some decreases which level off at 3-1/2.

This general trend perhaps reflects the fact that the child becomes more absorbed in the play as his attention span increases and that he shifts from one activity to another less frequently. In the early sessions the mother verbally stimulates more interaction, but as time goes on she does not need to do this so much and some 5 year old interactions are nearly silent.

In general there are positive correlations between sessions. Mother-child pairs who interact at a high level at one age tend to do so in later sessions, but the intersession correlations are generally lower for the child than for the mother.

Codability and comprehensibility

The West Harlem children increase significantly in codability from the first to the second session but from that point on the mean codability for all children's samples at all ages varies between 75% and 85%. The codability of the mother's utterances fluctuates randomly about 90% throughout.

The percentage of QQTC (questions about clarification of transient information) has been used as a rough index of the interruption of the communication between mother and child since utterances like "huh" or "What did you say" fall into this category. There are consistent decreases in this type of utterances with age for both mothers and children in all samples. This finding would be expected; the mother - child interaction seems to move more smoothly with each person better comprehending what the other is doing as the child grows older.

Questions and Statements

The analysis of questions and statements by the mother and the child has been analyzed into two factors first a so-called conversation quotient estimating the amount of the interaction that consists of exchange of statements (Mother's statements minus child's questions), second the estimate of the balance or imbalance of the questioning measured by the difference between the mother's questions and the child's. The conversation quotient generally varies from 25% to 45% and there are sample differences, but there is no clear evidence for age changes. If there are any at all, there is a tendency for the WH sample to reach about 40% by age four.

There is a consistent tendency for mothers to ask more questions than the child. The difference tends to decrease with age; twelve of the sixteen possible age comparisons are in that direction. (See Table 22) This suggests that with older children the questioning is more evenly balanced without one person quizzing the other.

Table 22

Mother-child differences in percentage of question

	2-1/2	3	3-1/2	4	4-1/2	5
WH	4.21	18.75	15.51	10.15	10.86	
CS		9.40		7.08		2.06
WS	21.09	12.02	8.26			

Behavior Requests

Another sign of imbalance in the mother-child interaction is the imbalance in the number of behavior requests. Mothers consistently make more behavior requests of the child than visa versa. Behavior requests generally decrease with age for both mothers and children. The mother-child difference perhaps decreases slightly, but the imbalance is nearly as great at 5 as at 2-1/2. The relatively sharp drop in children's behavior requests between WS sessions 3 and 3-1/2, that is matched by a similar drop in the WH data between 3-1/2 and 4,is suspicious since the change in coders occurred at that point. The failure to find a similar drop in the mother's behavior requests at that point, however argues against a coding change.

Content of Information

The utterances were distributed among four categories with sufficient frequency that they must be thought of as relative frequencies. The picture of age changes in these categories are very confusing, because all the changes that occur are non-linear. They go up or down at some ages and the reverse at other ages.

In general permanent information remains constant over age except among the Washington Square children where it drops in the middle of the three sessions. In West Harlem transient information and behavioral information show non-linear age changes that are mirror images of each other. Transient information increases then decreases while behavioral information does the reverse. Several of the changes are significant from one session to another. Among the Washington Square children, behavioral information remains constant but transient information mirror images the age changes in permanent information. Among Washington Square

mothers all categories remain constant over age.

None of these changes are readily interpretable, but since there are several significant age changes, they can hardly be attributed to random fluctuation. At present, however, they must be left uninterpreted.

Fantasy information remains relatively constant at a low level for all the samples. There is one significant difference between age 3 and 5 in the cross sectional sample, but the increase is not confirmed in the longitudinal data.

Mode of Information Exchange

Nine subcategories were defined for mode. The results on three of these categories, explanation, clarification, and labels, indicate significant and reliable trends which will be reported below. The means of the remaining categories were either based on too few observations to be reliable, or did not generate any significant and consistent differences.

Explanation

For all groups the amount of explanation increases with age. With WH this can be seen as a gradual increase but with WS it is a sudden increment. For the WS mothers, this jump (3% to 6%) takes place between ages 2-1/2 and 3, and is significant ($t = -3.30$, $p < .01$). This jump from (2% to 5%) for the WS children takes place between the ages of 3 and 3-1/2 and is also significant ($t = -3.24$, $p < .05$).

As noted earlier, the amount of explanation also increases in the Harlem sample, but not in a sudden increment as was the case with the WS sample. The gradual increase for the WH children is consistent but not significant while for the WH mothers the total change from 2-1/2 to 4-1/2 does prove to be significant ($t = -3.82$, $p < .05$).

The interpretation of a general increase does not seem difficult. Explanations are more complicated both to give and to understand than descriptions or labels, for example. The average syntactic complexity of explanations is high relative to many other categories. Thus the increase with age probably reflects the growth of the child's cognitive functioning and the mother's recognitiion of it. He asks for more explanation and she gives it, and he can also explain some things himself.

Labels

This interpretation of explanations is confirmed by the gradual reduction in the percentage of labeling with age. Only one significant difference was found among all the groups, but a consistent trend is evident in all of the groups. The single significant difference was

found in the WS child group where the amount of labeling from the 25% level at age 2-1/2 decreased to the 21% level at age 3 ($t = 2.43$, $p < .05$).

Although none of the other age to age differences are statistically significant, there is a consistent tendency from group to group to use less labeling as the child's age increases.

Clarification also decreased with age but it is due to the decreased QQTC which has been discussed under codability.

Responsiveness

The percentage of utterances which elicit a response shows a fluctuating pattern. In the longitudinal data in every sample there is a decrease in the percentage of utterances responded to and then a significant reversal of this trend, but the point at which the reversal occurs is 3 for the Washington Square sample and 3-1/2 for West Harlem. The cross sectional sample shows a significant increase between ages 3 and 4 for the mothers only.

Thus the results are difficult to interpret. A more careful analysis of the reasons for non-response is needed. If with increasing age the interaction becomes a less rapid-fire exchange of short utterances, then there would be more utterances embedded in short monologues and thus not responded to. On the other hand it would be expected that the child as he acquires interactional norms would become more responsive to those utterances that demanded a response.

The last expectation is clearly confirmed by the analysis of age trends in the responses to questions. The percentage of the mother's questions to which the child directly responded increases in all samples and many of the session-to-session shifts are significant. The effect is to equalize the responsiveness of the mothers and children from about age three on. There is also an increase in percentage of statements directly responded to.

On the other side there is a clear decline in the percentage of utterances responded to peripherally. This decline takes place between 2-1/2 and 3 and is significant in all samples. The no-information responses are more difficult to interpret. There are some significant session-to-session fluctuations, but they are not all in the same direction. The overall effect is no change. Mothers consistently give more no-information responses to the child's questions than vice versa, but the size of the difference does not change significantly.

One final relevant variable is the percentage of utterances ignored. This variable drops sharply as a result of a coding change which took place at a different point in the various samples and creates an artifactual age change and sample difference. On the other hand a drop would not be unexpected and does appear in the cross sectional data which are not so vulnerable to the change in coding.

Syntactic Complexity

Fortunately we can report that the grammatical complexity of the child's utterances increases steadily with age in all samples and the amount of change over any one year period is significant statistically If this age trend had failed to appear, we would certainly be suspicious of the validity of the index of syntactic complexity.

What is considerable more interesting is that the syntactic complexity of the mother's utterances also increases with the age of the child, but stays well below the complexity level found in her discussion with the adult interviewer after the first session. Her complexity leads the child's by 1-1/2 to 2 points. Table 23 shows the size of the mother-child difference at various ages in the various samples.

Table 23

The difference in complexity of the mother and child

Age	2-1/2	3	3-1/2	4	4-1/2	5
WH (long)	1.56	1.72	1.47	1.31	1.17	
CS		1.42		1.44		.93
WS	1.93	1.65	1.85			

There are two explanations for such a mother-child difference, depending on which factor is considered the hen and which the egg. On the one hand the mother's responses to the child's more complex grammar may be utterances that are themselves more complex. Explanations, for example, are more complicated than labeling utterances and thus a conversation involving explanations would be at a more complex level than one limited to labeling. The increase in complexity is not however merely a shift in frequencies from one VINEX category to another, because nearly every VINEX category itself shows an increase in grammatical complexity with age. Nevertheless the mother's complexity can be in response to the child's.

On the other hand the mother may be encouraging the child's linguistic development by constantly talking to him in a grammar that slightly, but not too much more complex than his own utterances. If this is true, the mother is certainly not consciously providing this sort of model, but she may monitor her own utterances to keep them clear and comprehensible to the child and thus in fact use a simpler grammar with the young child than the older one or when unconstrained by the age and maturity of the listener.

To illustrate how this modelling might work. Table 24 shows the WH mother's and child's utterances categorized according to how two independent ideas or kernel sentences are put together. These all represent double-based transformations and the frequency of each is calculated as a percentage of the total number of double-based transformations. Some of them, the simpler conjoining of two kernel sentences, decrease in relative frequency with age while the more complex types increase in relative frequency, but it is clear that none of the trends are simple linear ones. If we calculate for each percentage value for the mother, whether it is closer to the child's value at that age level or to his value six months later, we find that of the sixteen times when such a comparison can be made, her value is closer to the child's six months later ten times, in two cases the differences are identical and in four cases she is closer to the child's contemporaneous value. This suggests that in some sense she is leading the child, not only in average complexity but also in the frequency of various qualitative features of the grammar. The relative distribution of these transformations approximates their distribution in the child's utterances six months later. The data will permit a much more detailed analysis of the specific features of the child's grammar and the mother's grammar at different age levels, and the testing of the modelling hypothesis but this analysis has not been completed at the time of this report.

Table 24

Total Percentage of double-based transformations belonging in each category in West Harlem Mothers and Children

		AGE				
		2-1/2	3	3-1/2	4	4-1/2
Conjoining	C	64	42	32	27	28
	M	37	18	20	27	34
Subordinating	C	6	8	10	19	15
	M	12	15	16	17	10
Relativizing	C	3	23	21	21	21
	M	21	24	23	21	19
Embedding	C	28	27	38	33	36
	M	30	42	40	35	37

Summary

First the sensitivity of the VINEX measures and the syntactic measures have been established. They are capable of discriminating many features of the mother-child interaction and particularly in the mother-child comparisons and the age differences the variations revealed by the measures are sensible and expected.

Second, there is a general pattern of mother-child interaction. By far the biggest differences are in the relative frequency of different categories, e.g. requests versus commands, description versus explanation and the like. This points to an overall general pattern that is a sort of base line. The reasons for the general pattern is not at all clear, but must reflect common features of human interaction in our culture within the constraints of this particular environment.

Third, age differences are generally more prominent and more explicable than sample differences.

Fourth, therefore the relative paucity of sample differences can be taken to indicate the fact that they are not very marked. The absence of sample differences cannot be due to insensitivity of the measure, because the measure is sensitive to other expected differences. Of the differences between West Harlem and Washington Square, only two clearly suggest a cognitive deficit in the West Harlem children. One is the frequency of explanations, and the other is the syntactic complexity. The other sample differences are probably indicative of qualitative differences in the style of interaction but in no sense do they indicate that these Negro mother-child pairs are operating ineffectively. Even where the differences do carry this implication, they do not at all justify the common impression that the ghetto child's interactions with his mother and the middle class white child's interaction with his mother are different by a whole order of magnitude. Even when statistically significant, they appear to be only minor variations in a common human pattern of interaction shared by all members of our society.

The performance of children in standardized test situations

While the children's natural language in the playroom is a far richer source of data than their performance on standardized measures, the controlled stimuli of experimental situations provide information about cognitive development. Four measures were employed, a picture vocabulary test, the Wechsler Preschool Intelligence Test, the Berko measure of the childs grasp of morphological rules (Wugs Test) and a sentence imitation task.

Vocabulary

Of the language measures the Stanford-Binet Picture Vocabulary (Form L-M) was by far the easiest for the children. Sample differences between the West Harlem group and the Washington Square children at ages 2-1/2, 3, and 3-1/2 confirm the findings of other investigators--most West Harlem children perform less well than the Washington Sqaure children. The difference appears both in the number of children able to produce a scorable record, and also in the means and ranges of scores. (See Table 25)

However it is important to note that by the age of three similar numbers of children in the two samples respond to the test, suggesting that familiarity with a test situation increases the children's confidence and willingness to try. Also, at all three ages the mean score of the West Harlem children, although lower than the mean for the Washington Square sample, is high enough to pass the item on the Stanford-Binet at the appropriate age level.

Table 25

Stanford-Binet Picture Vocabulary

Age	Total N		Total Responding		Range of Correct Response		Mean		Passing Score
	WH	WS	WH	WS	WH	WS	WH	WS	
2-1/2	11	13	5	11	6-12	9-12	8.4	10.9	8
3	10	12	8	12	6-12	9-15	10.5	11.5	10
3-1/2	10	12	10	11	9-14	11-18	11.7	16.2	12*

*estimated--not administered on Binet at 3-1/2 years

The Wechsler Preschool Intelligence Test

The WPPSI was administered to the entire longitudinal Harlem sample at age 4-1/2. The range of I.Q.'s was wide, 63 to 128, with half the group within the normal range. For nearly all the children, performance scores were higher than verbal scores. All children testing within the normal range had been above the mean on the picture vocabulary at earlier ages.

The WPPSI is presently being administered to the Washington Square sample.

Berko Measure of Morphological Rules

The Berko Wugs measure was designed to study children's understanding of the morphological rules for forming plurals and possessive of nouns, as well as the third person present, the past and the progressive tenses of verbs.

Findings:

1. The major problem in using this measure with American Negro children is that many of the required inflections may be omitted in the non-standard Negro dialect as described by Loban, Labov and others. For example, it is impossible to tell whether the child's response of "rick" rather than "ricked" is an inability to form the past tense, or an encoding of the past in his own dialect.

2. The measure was very difficult for all children under four years of age. Berko did not use it with children under four and we did so only experimentally. In both the Washington Square and West Harlem samples at ages 3 and 3-1/2, 29 to 100 % of the children refused to respond to certain items, and there were very few sample differences in the frequency of this "no response" category.

3. Of the twenty five-year-olds in the cross sectional Harlem sample, sixteen children produced scorable records. The percentage of correct responses for these sixteen is quite different from the responses of the Berko pre-school group (composed of ten 4, 4-1/2, and 5 year-olds from the Harvard Preschool in Cambridge).

4. Children were most successful on forming the plural of Wug: 71% of our sample and 75% of Berko's giving a correct response. On all other plurals (tasses, glasses, heaf) the West Harlem group gave less than half as many correct responses as the Harvard group.

In the formation of the past tense the cross-sectional five-year-olds' correct responses were 30% to 66% lower than the Berko sample. This could be predicted in the light of findings on the Negro dialect. (Labov, 1968).

	Cross-sectional	Berko
glinged	12%	63%
ricked	48%	73%
melted	06%	72%
motted	0%	32%

One surprising reversal was that 12% of our sample formed the irregular past tense of a real word "ring", whereas none of the Berko sample gave a correct response. When required to form the third person present ending, (loodges), 18% of the West Harlem sample and 57% of Berko's sample passed the item.

Despite the overall differences in the number of errors the rank-order correlation between the difficulty of the items for our five-year-old group and for the Berko group of pre-schoolers was .74.

Analysis of errors of five-year-old responses

The errors the children made on this measure are quite consistent. When asked to form plurals, for example, they either repeated the stimulus word ("tass") or expressed the plural by a number ("two tass"). In forming the past tense, repetition of the stimulus word was the most frequent error: melt -- "melt"; rick -- "rick". Adding an auxiliary "did" would have been analogous to expressing the plural by the number "2", but nobody performed this way.

When the nonsense word was a verb the children sometimes described the action in ordinary English rather than inflecting the nonsense verb. This is not an illogical answer.

For example, the presentation of "zib" reads as follows:
"This is a man who knows how to zib. He is zibbing. He did the same thing yesterday. What did he do yesterday?
Yesterday he _____?" Answer: "Put a ball on his nose." This is not a bad description of the picture illustrating "zibbing."

This is a problem in constructing the measure. It is always difficult to present nonsense verbs pictorially. The verbs in picture vocabulary tests are always a source of difficulty. Unless the child really understands what he is expected to do , namely to inflect the nonsense word, these items are difficult. If he does understand the talk he probably can succeed but when he fails it is hard to know if he doesn't grasp the problem or has not acquired the rule.

Whatever the reason, many of the West Harlem five-year-olds do not demonstrate the knowledge of morphological inflections required for speaking standard English.

Our overall impression of the measure is that with further developments it might be a very useful diagnostic measure for analyzing the problem which some Negro children have with standard English. If items could be devised some of which reflected the non-standard dialect and others did not involve the dialect, then it might be possible to discriminate between grammatical immaturity and the speaking of non-standard English.

Imitation task

The sentence repetition task (See Appendix F) was presented to the five-year-old children, and analyzed by the method described by Osser, Wang, and Zaid (1969). We used only one of their analyses, the CSE (critical-structure error score). This is the number of sentences in which the subject made at least one error in the critical structure of the sentence being imitated. Since we administered only half as many sentences the scores are reduced to percentages. We also used their criteria for adjustment for non-standard dialect differences, since all the children in this analysis were American Negro. These characteristics of the non-standard dialect were taken from Loban's work (1966), and include: a) absence of third-person singular marker "s" on the present tense of verbs, b) omission of the verb "to be", c) omission of auxiliary verbs, and d) nonstandard use of verb forms. Osser also added a fifth characteristic from his own work, namely elision of the possessive marker "s". Actually it is only the latter rule which enters into this measure, and indeed 12 of the 18 children in our sample who passed this item deleted the "s" and responded "pulling the girl' hair". Like many of the Wug items, this example illustrates the ambiguity in selecting items measuring the language of American Negro children.

Our findings regarding LE children (children whose families have less education than our middle group) are very similar to the Osser, Wang, and Zaid findings. The mean critical error score for our 10 LE children, taking into account dialect differences, was 2.27 or 32%. The mean for the Osser group of 16 Negro lower class children was 45 errors or 34%. There was a difference between the two educational levels. The mean error for the lower educational level (LE) is 2.27 and for the ME group the mean error score is 1.11. Imitation of the reflexive showed considerably higher percentage of errors in the Osser sample than ours, probably because we scored the morphological error "hisself" for himself correct.

In summary, our investigations confirm those of Labov and others that the language of the young Negro child can be rich, complex and varied under optimal circumstances but that the standardized testing situation does always not tap the complex language of which the child is capable, not the language which he can comprehend. Obviously better measures and techniques must be developed if investigators are to understand adequately the cognitive and language development of the Negro child.

Whatever the reason, [illegible] demonstrate the knowledge of morphological inflections required [illegible] speaking standard En[illegible]

Chapter 4

Summary and Conclusions

When this study of adult-child interaction was first begun, its central objective was the development of methods for the description of interpersonal interaction as a technique for the investigation of cognitive socialization. The basic premises as outlined in Chapter 1, were, first that the socialization process must take place in the day-by-day interaction between the socialization agent and the child. Second, this interaction was assumed to be guided by the socializer's intuitive theory of human behavior, particular child behavior and development. Third, an effective method for recording this interaction was through the descriptions of an observer, who because he shared the intuitive theories of the mother and the child, and because ordinary language contains many words that reflect this intuitive theory, would produce a coded record of the interaction in his ordinary language description of it. Fourth, this ordinary language description would be in a sufficiently simple and standard grammar that it would be possible to write a computer program that could content analyze the observer's narration.

In the course of the investigation three methods for describing the adult-child interaction have been developed. These are the VINEX category system for coding the actual language of the adult and the child, a coding system for describing the non verbal behavior of the individual, and Interactional Language, for the use of an observer in narrating the adult-child interaction. These measures are described in Chapter 2 and coding manuals are available in appendices. The computerized analysis of the narration is well developed although no empirical results from the analysis are available at the present time. The essentials of the decoding program and the assignment of features to the words in the narration are described in Chapter 2 and in an appendix.

A method of syntactic analysis was adopted from the work of Sheldon M. Frank who has worked with us for several years. This syntactic analysis, while not one of the original objectives, has proved to be a valuable method which uncovers interesting information about the adult-child interaction and promises to allow the development of a descriptive grammar of the mother's language when talking to children of different ages. The analytic method as well as the scoring of syntactic complexity is described in Chapter 2 and a coding manual is available in an appendix.

In order to obtain material on which these methods could be tried and tested, we procured small samples of mother-child pairs, one from West Harlem, which included both a middle and a lower class subsample, and a white sample from upper middle-class homes in the neighborhood of Washington Square. About twenty of these mother-child pairs were observed repeatedly every six months, and other cross sectional samples at different ages provide data on mother-child interaction from 2-1/2 years of age to 5. All in all about 150 half hour interactions have been recorded, coded and analyzed.

Before reporting on these empirical findings, it is important to discuss the possible sources of error or misinterpretation. The primary objective of the research was the development of methods, not the actual description of the interaction of representative samples of children with their mothers. Furthermore the actual interaction observed was not expected to be a representative sample of the child's total pattern of interaction with his mother. The results are valuable nevertheless and do reflect important features of the adult-child interaction. At the same time the possible sources of error should be clearly stated.

First, the samples of families reached in both the Washington Square sample and the West Harlem sample are not representative of any well-defined population. In Washington Square the mothers were recruited through a play-ground association and surely represent mothers who are explicitly concerned with the welfare of their children. Other families in the area can be very different, as is apparent to anyone who has watched the passing parade from a bench in Washington Square park.

The sampling procedure for the West Harlem study was much more systematic but it is important to realize that many of the names originally obtained from birth records could not be located. Furthermore a certain family stability was necessary for participation in the study, namely the existence of a natural mother-child pair living together where the mother was willing to participate. The most extreme kinds of pathology that exist in Harlem were not represented any more than in Washington Square. In surveying the histories of a group of hospitalized adolescent boys from Harlem and the Bronx, it was apparent that none of them could have been participants in such a study as ours because they were in foster homes or living only spasmodically with the mother during the preschool period.

Furthermore any differences between the Washington Square and West Harlem samples cannot be attributed to any single factor. They differ in ethnic background, in social affluence, and in the kind of extrafamiliar environment in which they live. While children in both samples were exposed to urban crowding, dirt, and casual exposure to drunkeness, drug addiction and other pathology in the neighborhood, the Washington Square sample were certainly exposed to more desirable

housing and less crowding.

In regard to the sampling of the child's environment, it is clear that a half hour in a free play situation is not representative of the child's total socialization environment. In the home the child meets his father, his siblings, and his mother is necessarily busy much of the time and unable to give him her undivided attention as she can in the playroom.

While there is no question about the need for observational records in the home, the importance of observational records in a standardized situation must not be underestimated. The behavior of the mother and the child in the playroom stems from their stable interaction patterns. The mother may be trying to put her best foot forward, but what she thinks is her best foot depends upon her normal pattern of interaction. Furthermore no mother can maintain a totally false pretense over a half hour in the presence of a child who is certainly not trying to put his best foot forward. The loss in representativeness is counterbalanced furthermore by the fact that the physical stimuli of the play room do not vary from one pair to another and we have seen that these stimuli affect the variables being observed. The validity of findings do not depend on replication of the natural environment in all details; the intelligence test, for example, although not a natural environment, yields valuable information.

What is ultimately needed is psychological ecological study of children from various environments and the problems of such ecological investigations have not been solved -- certainly we have not solved them. To assume that any observation that is not in the natural setting is worthless badly distorts what ecological studies contain. In biological ecology for example, some observations are necessarily made in the natural environment but many of the studies of the flora and fauna that contribute to an adequate picture of the ecological community are made in greenhouses, aquaria and even in the test tube.

With these considerations about what this study of mother-child interaction can and cannot provide, we turn to a consideration of the empirical findings.

The first empirical finding is the general pattern of mother-child interaction, the base line around which different samples and different age groups vary. This pattern is partly a result of various interactional norms in our society, and partly a function of the specific situation, namely a play room in which the child could play with anything he wanted and in which the mother did not have any specific role except to play with the child.

Under these conditions we find that the interaction contains both conversational exchange of statements as well as question and answer sequences, about one third of the former and two thirds of the latter, with the mother leading the child somewhat in questioning. In general, however, the burden of the interaction is divided about evenly between the mother and child. The mother requests more behavior of the child than visa versa. The interaction is concerned with permanent information about one third of the time and is heavily weighted with descriptive statements and specifications. Labeling and explanation are less common. Requests for clarification are not uncommon, about 15 to 20%, and seem to reflect interruptions in the communication between mother and the child.

The interaction is, generally speaking, a responsive one. Some utterances are response demanding like questions; others may elicit a response but do not require one; still others, like statements in the middle of a speech, cannot be responded to without interrupting the speaker. About 70% of the questions in these interactions are responded to by the other participant and only 1% or so are clearly ignored. We have not devised a coding for interruptions, but in other respects the interaction seems to conform to these general norms of human interaction. Without observing more varieties of interactions we cannot know what part of this pattern is a result of its being a mother and a child, what part its being a play situation, and what parts are generally invariant in human interaction. It is clear however that the interaction was not primarily a didactic lecture, an oral examination or a passive mother watching her child play, or two people each minding his own business. It was a give and take interaction, not precisely balanced, but at the same time not markedly one-sided.

The changes in the interaction with age seem partly attributable to the child's cognitive development, increase of explanations, and increased grammatical complexity for example. Other age changes can be seen as the child's acquisition of the norms of interaction, e.g. codability of utterances decrease of QQTC utterances, increased frequency of response by the child. Still a third developmental trend is toward the equalization of the two parties in the interaction which would generally exist if both participants were the same age. The equalization is of course not attained by the time the child is five.

Against the general pattern, the age differences are like waves in a stream; they are significant, but the interaction pattern does not change drastically or dramatically from one age to another as it might if there were clear developmental stages. We suspect incidentally,

that there may be such a dramatic shift in mother child interaction between the time when the baby is clearly unable to decode the content of speech, age 6 months perhaps, and the time when the mother expects him to understand what she says to him and consequently monitors her language to be as simple and comprehensible as it can be. But over the age range studied here the age changes are not sharp clear steps, they are trends.

On this background the differences between the West Harlem and the Washington Square sample, as well as the difference between the lower and middle class Harlem samples appears very small. Sample differences are generally less significant than age differences. There are some significant sample differences consistent across ages in conversational quotient, in amount of clarification, in explanations, and in syntactic complexity. But we have no data to confirm beliefs that the ghetto child does not have meaningful conversation with his parents or that he is presented with a totally inadequate language model that makes him incomprehensible to a nursery-school or kindergarten teacher.

Such beliefs are not supported by this study on two grounds. First the individual differences among the Harlem families are large, probably larger that in most non-ghetto neighborhoods. This diversity is a result of the fact that Harlem is a ghetto which Negro families find it difficult to escape from because of the restriction of their free choice in housing. Thus it compresses into one area many diverse kinds of families that would in freer circumstances spread out into many different types of neighborhoods. There are common features to the Harlem environment that probably have some homogenizing influences on family patterns, but the lack of free self selection makes Harlem more diverse than many other neighborhoods.

Secondly the commonly held stereotypes of the family interaction of Harlem children were not supported by any of our empirical data. No single mother-child pair fits the stereotype. We do find differences between Harlem and Washington Square, but they are again variations in a general theme, not different melodies.

What makes this finding particularly interesting is that the West Harlem children in more standardized testing situations did respond very differently from the Washington Square children. Thus we replicated the finding of an apparent difference in cognitive functioning in test situations. What leads to a distrust of the validity of these test differences is the marked change in the apparent cognitive functioning of the West Harlem children when they went from the test to the playroom.

The importance of the test scores should not be understated. Tests are a standard feature of our society and they represent one of the facts of life that children must adapt to, even if they are unfair in

individual cases. This finding leads therefore to the need for further investigations of the sources of these tested differences. Is it merely an example of the fact that psychological examiners do not always measure the child's maximum potential, or is there something in the Harlem child's interactions with his environment that leads him to be more vulnerable to the standardized testing situation? Both are possibilities that must be resolved empirically.

In summation this study has certainly raised more questions than it has resolved. Its major contribution is that it has lead to new methods for describing interpersonal interaction, which can be useful in further studies to help answer the questions that our empirical data raises.

APPENDIX A

MANUAL FOR MEASURING SYNTACTIC COMPLEXITY

Sheldon M. Frank, M.D.
67 Sky Meadow Place
Elmsford, N.Y. 10523

Milton S. Seegmiller
Cornell University
Ithaca, N.Y. 14850

December, 1969

Preface

Table of Contents

MANUAL FOR MEASURING SYNTACTIC COMPLEXITY

Sheldon M. Frank and Milton S. Seegmiller
Cornell University
Center for Research in Education

PREFACE

The germ of the analysis scheme described in this manual began while the undersigned was collaborating with Harry Osser at Johns Hopkins University Medical School in 1965-1967. To him, and to Alfred and Clara Baldwin, with whom psycholinguistic collaboration was further pursued from 1967 through the present time, is owed the greatest appreciation both for their eager and loyal encouragement of the project of formulating such a system and for their imaginative and elegant work in devising experimental frameworks in which such a system could play a role. To none of these three, however, should be ascribed any of the possible linguistic shortcomings of the system. Neither should these be blamed on Milton S. Seegmiller, who has joined the collaboration in 1968; many of his keen linguistic suggestions will not be able to be employed until the series of experiments begun with the present system are finished. His contributions in making modifications were possible and in clarifying and helping to finalize a large set of rules and customs (which had sometimes grown in a most amorphous way) were of great importance. Finally much gratitude must go to the research teams of Osser at Johns Hopkins and of the Baldwins first at New York University and now at Cornell; many indeed were the suggestions and modifications proposed by several workers among all three groups. Special gratitude must go to Jan Drucker, Susan Feldman, Zena Farbstein, Mary Jane Murphy, Bonni Seegmiller and Marilyn Wangh.

SYNTACTIC ANALYSIS SYSTEM

I. INTRODUCTION

The notions of syntax and grammar underlying the present work are mainly those of Chomsky (1957, 1965). The approach of Roberts (1964) in translating these notions into a practical transformational-generative grammar of English has been an important model. To other workers in both linguistics and psychology must go credit for supplying us ideas for various discrete formulations (e.g. Menyuk for the inversion transformations.)

We have attempted to make our grammar as inclusive as possible for the speech of preschool and kindergarten children; it does not pretend to be a complete grammar of adult English. Furthermore, as many workers in psycho-linguistics, we too must acknowledge a few years lag behind the most recent modifications in the swiftly developing field of transformational-generative grammar. The option thus chosen - of not frequently modifying the system to keep abreast of innovations - carried for us, beside the onus of old-fashionedness, the advantage of facilitating longitudinal studies by retaining comparability between successive sets of data. Our system, we think, has been useful thus far as a research tool (see Bibliography, Section II) which of necessity must be continually subjected to review and revision, even though the changes may sometimes have to wait for their incorporation into the system.

The basis for obtaining a quantitative measure of the syntactic complexity of a set of utterances is the idea that arbitrary yet consistent numerical values can be assigned to certain basic features of sentence structure such that the scores thus obtained can be used 1) in comparing relative complexity of different utterances within the output of a speaker or 2) that total and/or mean values based on the scores can be used to compare relative complexity among different speakers or for one speaker in different situations.

The values assigned are arbitrary in the sense that they cannot be proven to express any objective reality, i.e. there is no universal psychologically, linguistically, or neurologically based scoring system which requires that a noun phrase like "the man" be assigned a complexity score of one unit. However, it seems completely reasonable to assert for example that a set of noun phrases such as "the man," "a boy", "this car," etc., are all of more or less equal complexity; or that when a white middle-class female adult says "the man," the complexity of her utterance is approximately the same as that of a black working class five-year-old boy when he says the same thing. The only other assertion made by the Syntactic Complexity Measure is that certain elements and operations in syntax are roughly equivalent to each other, e.g. that a noun phrase, an adverb of manner ("quickly") and transposition of an element within a sentence should each be

assigned the same complexity score. This is a more difficult assertion to defend and will have to be accepted as a postulate for the present.

The Syntactic Complexity Measure is made up of two sets of rules: The Base, or Phrase-Structure rules (P-rules) and the Transformational rules (T-rules). The P-rules derive or "generate" (and, therefore, may be used to analyze)[3] relatively simple kernel sentences, to which one or more T-rules may later be applied to produce transformed sentences. Accordingly, a sentence will be called Transformed if any T-rule has been applied to it, and a Kernel sentence otherwise.

II. THE BASE COMPONENT

A. General

The Base Component contains a number of Phrase-Structure rules (P-rules) of the form A --> B+C, which may be read as "A is to be rewritten as B plus C." The initial P-rule is

(1) S --> NP + VP

which states that a sentence (normally) contains a Noun Phrase and a Verb Phrase. Since this is the basic structure of a sentence, we will arbitrarily assign one unit of complexity to each of these elements, i. e. NP and VP. Therefore, the sentence $\overset{np}{\underset{\text{"John runs"}}{}}\ \overset{vp}{}$ is scored two points.

Every NP and every VP contains certain obligatory elements and may contain other optional elements. All obligatory elements taken together constitute the essential part of the NP or VP and will be awarded only one unit no matter how many words may be included. Thus, "the ball," "democracy," "runs", "played with some toys" are each scored as one unit.

Each optional element will be awarded one additional point, e.g. "the blue ball" will be scored two units.

---

[3] Transformational-Generative Grammar, as the name implies, is concerned with generating sentences, beginning with the most basic elements (S, N P, VP) and proceeding via a set of rules to the surface structure, i.e. the way the sentence is spoken or written. Obviously, the opposite procedure will be used in scoring syntactic complexity, that is, to begin with the surface structure and work backwards through the rules to discover the underlying structure of the sentence. However, in the following sections, when the rules are first presented, we will use terms like "selecting elements" or "producing phrases" as though sentences were being generated rather than analyzed.

B. The VP

The essential element of a VP is the verb, and any verb present in a sentence will therefore be worth one unit. However, since certain types of verbs must be accompanied by obligatory elements, naturally they will not be assigned any additional units. For example, "hits the ball" is a verb phrase with two obligatory elements: a transitive verb (VT) and its object, and is scored one unit. Note that an NP which functions as the object of a verb is not awarded an additional unit.

The P-rule for VP is:

(2) VP --> Aux + { a) be + { (i) adv - prep.p / (i.i.) substantive / (iii) adv - p } + [Adv.] / b) verbal }

Aux is an obligatory element containing the tense marker

where

{ } denotes obligatory selection of an element within the brackets

adv. = adverb
prep.p = prepositional phrase
substantive = a noun or adjective
p. = place
[ ] signifies an optional element

verbal will be explained below

Rule (2) is an abbreviation of all of the following rules:
(2) (a) (i) VP --> aux + be + adv - prep.p (e.g. "is with his mother")
(2) (a) (ii) VP --> aux + be + substantive ("is a house", "is blue")
(2) (a) (iii) VP --> aux + be + adv-p ("is here")
(2) (b) VP --> Aux + verbal.

The element "verbal" specifies the various possible types of verbs other than "be", and may be expanded by Rule (3):

(3) verbal --> { (a) VI; (b) VT + NP; (c) VB + { (i) NP; (ii) [int] + adj }; (d) Vs + [int] + adj; (e) Vh + NP }

where

VI = intransitive verb
VT = transitive verb
Vb --> become, etc.
int = intensifier (very, etc.)
adj = adjective
Vs --> seem, appear, feel, etc.
Vh --> have, etc.

For example:

(3) (a) VI: cry, go, disappear.
(3) (b) VT + NP: hit the ball, eat lunch.
(3) (c) (i) Vh + NP: become president
(3) (c) (ii) VB + [int] + adj: become [very] silly
(3) (d) (Vs + [int] + adj: seem [very] tired
(3) (e) Vh + NP : have a headache

There is one important point that should be kept in mind with regard to VP's: it is the <u>verb</u> that is the essential element, not the NP or the adjective, and in those cases (which are not uncommon in children's speech) where the verb is present but an "obligatory" NP or adjective is not (e.g. "I am." "he hit,"), one unit will be given for a VP anyway. However, the converse is <u>not</u> true; in sentences like "I tired" or "This a ball", where an adj. or an NP is present but the verb is missing, no units will be scored for the VP. In other words, we are distinguishing between <u>essential</u> (that which is required if an element is to be awarded units of complexity) and <u>obligatory</u> (that which is necessary for grammaticality).

Rules (4) through (7) expand the element <u>Aux</u>:

(4) Aux --> tense + [M] + [have + part] + [be + ing] + [aux-s + to]

(5) tense --> present, past.

(6) M --> can, will, may, shall, must.

(7) aux-s --> have, be supposed, be going, like, etc.

Note that tense is the only obligatory element of Aux. This is the element that, by means of a set of rules that don't concern us, provides the correct form of the verb to agree with the subject, e.g. "I am," "he is", "John hits", "they played", etc.

The selection of one of the optional modal words (M) provides for sentences like "I will go", "I may go," "I shall go", "I must go". Could, would, might, and should are also M's and, for our purposes, can be considered to be derived from can, will, may, and shall plus the tense element past. This can be stated in the form of rule (8):

(8) past + { can / will / may / shall } --> { could / would / might / should }

The category Aux-s contain a specialized group of words which commonly occur in structures of the type Aux_s + to + V (V will be used to signify "be" and verbals), e.g. "want to go", "have to eat", "is going to play", "are supposed to sing," "try to see" etc.

The two remaining (optional) members of Aux ([have + Part] , [be + ing]) provide for constructions which are often called participial. The elements part and ing form the past and present participles, respectively, of any verb which immediately follows. For example, if we select [have + part] from the Aux and VI from the verbal, we will have .. have + part + VI ... which corresponds to "have gone," "has eaten," etc. Similiarly, be + ing + Vb + NP might yield "is becoming a bore." If both [have + Part] and [be + ing] are selected, we might get "has been playing" or "have been talking."

The final VP rule is:

(9) Adv --> Adv-t, adv - p, adv-m, adv-freq.

This rule states that Adv may be an adverb of time (then, now, tomorrow, last year), of place (here, there, in the bathtub, at the store), of manner (quickly, with disdain), or of frequency (often, seldom, once a year, every Friday).

The following illustrate some possible types of VP's and the number of units they would receive.

Scoring of VP only

a. I/go (VI) 0 optional units, 1 total unit
b. I/may go. (M-VI) 1 optional unit, 2 total units
c. I/may have gone. (M-have+part. - VI) 2 optional units, 3 total
d. I/am supposed to go (Aux_s + to - VI) 1 optional, 2 total.
e. He/hits him with his hand (VT + NP - adv-m) 1 optional, 2 total.

f. I/am very angry (be-int.-adj) 1 optional, 2 total.
g. They/wander aimlessly everyday in the park. (VI - adv-m-adv-freq-adv-p) 3 optional, 4 total.

C. THE NP

The rewrite rules for NP:

(10) NP --> { (a) proper noun / (b) Det + N / (c) Pronoun } + [adj - prep.p.]

(11) Pronoun --> { Pron - pers. / Pron - indef / Demon. }

(12) Det --> [pre-art.] + { Demon. / Pron - poss. / Art } + [number] + [adj]

where

pre-art	-->	some of, several of, a few of ...
demon	-->	this, that, these, those
Pron-poss	-->	his, my, their...
art	-->	the, a, some
number	-->	one, two ....
adj	-->	blue, big, conceited,...

Rule (10) states that an NP must contain either a noun or a pronoun. These are the essential elements of an NP, and will be awarded one unit of complexity when present in any NP functioning as the subject of a sentence. (As noted earlier, this scoring procedure does not apply to NP's functioning as objects of verbals. However, the rules themselves apply to all NP's, and any optional elements within an object NP will be awarded additional units).

Det. is an obligatory, but not an essential, member of an NP containing a common noun, and will therefore not be scored separately, e.g. "Ball is blue" and "The ball is blue" will both be awarded two points, since both contain the essential parts of the NP and VP.

The optional element adj - prep.p provides for NP's like "The man from California" and "anyone with legs." An adj - prep.p will be scored one additional unit.

A det. must contain either a demonstrative, a possessive pronoun, or an article, and may also contain any one or more of three optional elements (which will be awarded one additional unit each); a pre-article, a number, or an adjective. Thus, the following are all possible NP's:

(a) John (proper noun) 0 optional units, 1 total unit
(b) The bartender (Det + N) 0 optional, 1 total
(c) Those men (Det + N) 0 optional, 1 total
(d) Everyone (Pron - indefinite) 0 optional, 1 total
(e) He (Pron-personal) 0 Optional, 1 total
(f) John of Runymede (Proper noun + adj-prep.p) 1 opt.,2 total
(g) Several of those ninety-two ugly ducklings
(preart. + Demon + number + adj + N) 3 opt., 4 total

D. SUMMARY

The entire Base Component can now be condensed into the following set of rules:

(1) S --> NP + VP

(2) VP --> Aux + { (a) be + { (i) adv - prep.p / (ii) substantive / (iii) adv - p } + [adv.] / (b) verbal }

(3) Verbal --> { (a) VI / (b) VT + NP / (c) Vb + { (i) NP / (ii) [int] + adj } / (d) V_s + [int] + adj / (e) Vh + NP }

(4) Aux --> tense + [M] + [have + Part] + [be + ing] + [aux_s + to]

(5) tense --> present, past.
(6) M --> can, will, may, shall, must
(7) Aux_s --> have, be supposed, be going, like, ...

(8) past + { can, will, may, shall } --> { could, would, might, should }

(9) Adv --> adv-t, adv-p, adv-m, adv-freq.

(10) NP --> { (a) Proper noun; (b) Det + N; (c) Pronoun } + [adj-prep.p.]

(11) Pronoun --> { Pron-pers, Pron-indef, Demon }

(12) Det --> [pre-art] + { (a) Demon; (b) Pron-poss; (c) Art } + [number] + [adj]

To illustrate the method of scoring Kernel sentences, a set of sentences has been selected from the transcript of one of our subjects. (K.D., age 4). The number in the left-hand column locates the sentence within the transcript. M: indicates that the mother is speaking, C: that the child is. Explanations and comments are provided when necessary.

122.	M:	That's wrong again. (Pron-demon + be + adj + [adv]) (contractions are not scored, that is = that's).	2 + 1 = 3 units (2 + 1 indicates 2 obligatory elements + 1 optional element)
123.	C:	That's wrong.	2 units
124:	M:	This is the truck.	2 units
126.	M:	A gas truck. (art - [adj] - N)	1 + 1 = 2 units
127.	C:	Truck.	1 unit
128.	M:	That's part of it. (pron-demon + be + N + adj-prep.p)	2 + 1 = 3 units
131.	C:	This go here. (pron-demon + VI + adv-p)	2 + 1 = 3 units

Since tense is not an essential element of the VP, it is not scored separately. This has the effect of eliminating any bias in the scoring of dialects where the third person singular ending - s is commonly omitted.

"Go" is a VI, therefore "here" is optional. If the verb were "be", "here" would not receive + 1.

212. C: We gotta do it on this one — 2 + 1 + 1s = 4 units
(pron-pos + aux_s + to + VT + pron-pos - adv-p) (1 = 4)
"Gotta" is taken to be a colloquial form of "have to" or "have got to". Aux-s's are indicated separately by an "s" following the number of units, thus, the scoring of this sentence indicates 2 units for "we do it," 1 unit for "on this one" (=adv-p) and 1s unit for "gotta".

216. C: I found one. — 2 units
Here, "one" functions as a noun, not as a numeral preceeding a noun.

217. M: That's a T. — 2 units
218. C: A T. — 1 unit
221. C: Here. — + 1
If no matrix (i.e. NP or VP) is present, optional units are still preceeded by a plus sign.

III. The Transformational Component

A. General

Transformational rules (T-rules) transform Kernel sentences into more complex sentences by means of one or more of the following operations:

1. Addition of an element
2. Subtraction of an element[4]
3. Transposition of an element
4. Changes in intonation (voice inflection).

For example, a T-rule which transposes an element might be written

---

[4] The question of subtraction or deletion of elements is currently being debated among transformational grammarians. By including "subtraction" among the transformational operations, we are making a purely pragmatic decision and are in no way taking a stand in the debate. In fact, it was decided to eliminate a separate deletion transformation from the measure since it is difficult to demonstrate that a deletion transformation has ever been applied. In other words, the coder is normally not able to decide if an element was present and has been deleted or if it was never present at all.

(13) A + B + C --> C + A + B

where the elements on the left side of the arrow comprise the "structural description" and describe what the sentence (or part of it) must look like before the rule can be applied; and the right side describes the "structural change," or what it looks like after the rule has been applied. A T-rule which both adds and subtracts an element might take the following form:

(14) A + B + C --> A + C + D

and a T-rule which changes intonation might make the following change:

(15) You know him. --> You know him?

where question intonation is indicated by the question mark.

In measuring the complexity of transformed sentences, one unit is scored for each of the four operations which is carried out. Thus, in scoring example (13), the name or number designating the particular T-rule (these are explained below) is entered on the scoring sheet with a "1" along side it to indicate one unit or operation. Example (14) would receive two units, example (15) one unit.

In addition to the above types of transformations, which are called "single-base T's", there is another type, called "double-base T's", which combines two separate Kernels into a single sentence. These rules take the following form:

(16) matrix: A + B (e.g. John goes)
insert: A + B (e.g. Tom stays)
result: A + B + C + A + B (John goes and Tom stays)

The same four operations (add, subtract, transpose, intonation) apply to double-base T's, but additional points must be scored for the "insert," i.e. the second Kernel sentence. Thus, the double-base T in example (16) would receive 2 units for the insert and 1 unit for the addition of the element "C", for a total of 3 units. The total score for (16) including the matrix, would be 5 units, (2 for the matrix and 3 for the T). If the insert kernel contains optional elements such as adjectives, adverbs, intensifiers, etc., optional units will be scored just as if these elements appeared in the matrix, except that they are entered in a separate column on the coding sheet. This will be illustrated in Section V.

B. Single Base Transformations

In the list which follows, the number of units indicated for each transformation includes only the four transformational operations; addition, subtraction, transposition, and intonation. Units for the

matrix are not included. It should be remembered that the complexity scores for these transformations are based on the complete operation of the transformation according to the rules presented here. Occasionally sentences are encountered which do not quite conform to the rule. In such cases, when the transformation is entered on the coding sheet, an asterisk is placed next to it and an appropriate number of units is added or subtracted. For examples, see Section V B.

The symbols W, X, Y, Z are used to designate elements which have no function in the operation of the transformation, and may stand for any element of any length or no element at all, unless otherwise specified. Thus, "X + Y" in a structural description indicates that 0, 1, or 2 elements may be present, and that the elements may be of any length, from part of a word to an entire sentence. However, restrictions such as "Y ≠ 0" (which states that Y cannot be a "zero" element, i.e. it must stand for something) or "X = NP, VP, adj, adv" (i.e. X must be one of these elements) may be placed on X and/or Y.

The four operations addition, subtraction, transposition, and inflection (intonation) are indicated here as ad, sub, tr, infl, respectively. The symbol "V" includes both "be" and "verbal". "Comp" is a "dummy" symbol indicating that an element (usually an insert kernel) will replace it.

At least one example is provided for each T-rule. In addition, supplementary examples have been provided for some rules for the purpose of illustrating points which have presented some difficulty for coders in the past or which are not obvious from the rule itself. For the most part, these are from transcripts of actual interactions and interviews, and, therefore, often reflec colloquial, rather than standard, usage. Comments and notes on both the rules and the examples are provided where appropriate.

SINGLE BASE TRANSFORMATIONS

1. T-adv-sm: X + Y --> X + Adv - sm + Y
   where adv-sm = an adverbial acting as a sentence modifier, e.g. maybe, though, etc.
   e.g. We didn't lose any of our money --> Fortunately, we didn't lose any of our money.
   Operations: Ad
   Units: 1
   Note: The adv-sm is not restricted to the initial position, e.g. This one, though, is for him.

2. T-affirm: a) Z + X + Y --> Z + X + A + Y
   (where A = primary stress for preceding word and X ≠ 0 or verbal)

   b) NP + tense + verbal --> NP + tense + A + verbal
   (requires T-do)
   e.g. a) I have spoken --> I HAVE spoken

b) I went --> I + tense + A + go (--> I DID go, after T-do)
Operations: ad
Units: 1

3. T-conj-P: X + P + Y --> X + P + conj + P + Y
where P = NP, VP or any subelements thereof and conj= and, or, but ...
e.g. We eat fish today --> We eat fish and cheese today.
Operations: ad (conj), ad (P)
Units: 2
Note: In some cases, T-con-P may be used with vocatives, e.g.

OK, Eric and Lisa, come on.

Otherwise, T-conj-P adds only NP, VP, or subelements thereof (i.e. matrix elements) not elements added by other T-rules. For example in

They're shorter and cleaner,

T-compar-a must be scored twice, since T-conj-p adds only "and clear," not "-er"; which is added by T-compar-a. Similarly, optional elements accompanying the element which is added are scored as optional unit in the matrix. E.G. in

I have a black cat and a brown dog.

T-conj-p adds "and a dog", and the adjectives "black" and "brown" are scored + 2 in the matrix. Compare the following:

I have a brown and black dog

where T-conj-p adds "and black" and only one optional element ("brown") is scored in the matrix.
Note that sentences like

Look around and come back.

(i.e. conjoined imperatives) are scored T-conj-K, not T-conj-P.

4. T-compara-a: a) X + adj + Y --> $\left\{\begin{array}{c} X + adj + er + Y \\ \text{or} \\ X + more + adj + Y \end{array}\right\}$

e.g. I am big --> I am big + er or
I am beautiful --> I am + more + beautiful
Operations: ad
Units: 1
Note: The presence of a comparative signals the presence of an optional adjective except when it is a predicate adjective after "be". It is important to

distinguish between "more" as a comparative and "more" as an adjective. That is, "more" is a comparative in "more beautiful" but an adjective in "more people." In the sentence

> You're becoming more acquainted with it and you're remembering more.

The first more is a comparative, the second is not. Similarly, "better" may be the comparative of "good" and "well", as in

> This one is good, but that one is better.
> I didn't feel well, but not I feel better.

or it may be an adverb, as in

> I think I'd better put it together.

5. T-do: X + tense + Y --> ... X + do + tense + Y, ..
where X ≠ be, verbal, M, have
e.g. (past) + John go there --> did John go there
Operations: ad
Units: 1
Note: The verb "do" has three major functions in English:

1. As a function word in questions and negatives, e.g. "Did you eat?", "I didn't eat."
2. As an emphatic word, as in "I did eat."
3. To replace another verb which one doesn't want to repeat as in
"You ate already?"
"Yes, I did."

For the purpose of this measure, T-do applies only to functions 1 & 2 -- i.e. only when another verb is present. In all other cases, "do" is assumed to be the main verb, and is scored as part of the matrix, as in the following:

> Yes, they do.
> Do it like you do at home.
> Didn't you?

In the following sentence, "do" is used as both the main verb and an emphatic word, and T-do is scored only once:

> Yes, he does do things like that.

6. T-Go (come) This transformation applies to the use of "Go" (and also Come) as an auxiliary to the main very of the sentence. There are some restrictions on its occurrence, but these have not been worked out in detail:

NP + VP --> NP + Go + VP

e.g. He is going to work for them -- He is going to go work for them.
Operations: ad
Units: 1

7. T-imper-a:

$$X + You + tense + \left\{\begin{matrix}verbal\\be\end{matrix}\right\} + Y \rightarrow$$

$$X + \left\{\begin{matrix}verbal\\be\end{matrix}\right\} + Y + (imper.\ inflection)$$

e.g. You go fast --> Go fast !
Operations: sub, infl
Units: 2
Note: In certain cases, "you" may not be deleted by T-imper-a. Exchanges like the following are often found in dialog:

Pick it up.
No, you pick it up.

The second sentence is scored as though "you" had been deleted, i.e. T-imper-a gets 2 points. Sentences followed by question marks usually cannot be imperatives, e.g.

See?

The exception to this is the case where a tag question has been added, as in the following:

Pick it up, OK?
Help me, all right?

8. T-imper-b:

$$X + we + tense + \left\{\begin{matrix}be\\verbal\end{matrix}\right\} + Y \rightarrow$$

$$X + let's + \left\{\begin{matrix}be\\verbal\end{matrix}\right\} + Y + (imper\ inflection)$$

e.g. We go fast --> Let's go fast !
Operations: sub, infl, ad
Units: 3

9. T-imper-c: $X + \begin{Bmatrix} we \\ I \end{Bmatrix} + tense + \begin{Bmatrix} be \\ verbal \end{Bmatrix} + Y \rightarrow$

$X + let + \begin{Bmatrix} we \\ I \end{Bmatrix} + tense + \begin{Bmatrix} be \\ verbal \end{Bmatrix} + Y$

(requires T-obj)
e.g. I go fast --> (after T-obj) Let me go fast!
We go fast --> (after T-obj) Let us go fast!
Units: 4

10. T-imperfect: NP + V + X --> NP + USED + TO + V + X -
e.g. He did that --> He used to do that
Operations: ad, ad
Units: 2
Note: Compare the following, where "used to" is not an imperfect:
You're used to him.
(V = to be used to)

11. T-indir obj: $X + VT + Y \rightarrow X + VT + NP_{io} + Y$
e.g. We told the story --> We told our children the story.
Operations: ad
Units: 1
Some other indir. objs:
I'm going to bid you goodbye.
Tell me.
Give me one.

12. T-interj: X + Y --> X + interj + Y
e.g. It's 10 o'clock --> Gee whiz, it's 10 o'clock.
Operations: ad
Units: 1
Note: The element "interj" is not restricted to any fixed position in the sentence --- it may appear anywhere. The following sentences illustrate several common interjections.

Hello.
Pardon.
Please.
Oh - oh.
No.
No, I didn't do it.
Yes.

Yes, I did.
OK.
OK, I'll help you.
All right, I'll help you.

} "OK is always an interjection; all right" is only when it accompanies something else. When used alone, "All right." is scored as a Kernel sentence with no verb. See section IV.

You fixed it, right? (when used by itself, "right" is 1 in the matrix as an adjective (i.e. assuming "that is right"), not an interjection.

13. T-invers: NP + VP --> conj or introd + NP + VP
where introd --> well, so
e.g. The man sees him -- So, the man sees him; but the man sees him; And the man sees him, etc.
Operations: ad
Units: 1

14. T-neg: a) $X + \begin{Bmatrix} have \\ M \\ be \end{Bmatrix} + Y \dashrightarrow + \begin{Bmatrix} have \\ M \\ be \end{Bmatrix} + not + Y$

b) X + tense + Y --> X + tense + not + Y (requires T-do) where neither X nor Y contains have, M, or be.

e.g. a) The boy has seen the world --> The boy has not seen the world.
b) The boy + past + eat supper (= the boy ate supper) --> The boy + past + not + eat supper (--> The boy didn't eat supper, after T-do)
Operations: ad
Units: 1
Note: That T-neg adds only the element "not" (which may also appear contracted as " ____n't"). Therefore, many sentences which must be considered semantic negatives cannot be formed by this transformation.* The following are some examples with the correct scoring indicated:

*
A comprehensive discussion of the various types of negatives, and a formulation of the necessary rules, may be found in: E.S. Klima, "Negation in English," in J.A. Fodor and J.J. Katz, "The Structure of Language: Readings in the Philosophy of Language," Englewood Cliffs, Prentice-Hall, 1964.

No, people fly in the water. (T-interj)
He never used language. (+ 1 in the matrix for an adverb)
No matter? (+1 in the matrix for an adjective).
Nothing is in here. ("Nothing" is a noun. The matrix = 2 units, with no optional units)

Often a sentence will contain both "not" inserted by T-neg and one of these other types of negatives, e.g.: I can't put nothing back on.
No, it's not this one.
The police wouldn't come for nothing.
I can't find no more.

In exceptional cases, T-neg may be applied when no verb is present:
Not you !
Not this one.

"Ain't" is a possible variant of T-neg.

15. T-obj: VT or VH or prep + personal pronoun --> VT or Vh or prep + personal pronoun + OM
e.g. They saw he --> they saw he + m (he + m --> him, by phonological rule)
Operations: ad
Units: 1

16. T-passive-a: NP_1 + Aux + VT + NP_2 --> NP_2 + Aux + be + part. + VT
e.g. The mouse killed the fly --> The fly was killed.
Operations: tr, ad, ad, sub
Units: 4
(See notes following T-passive-b).

17. T-passive-b: NP_1 + Aux + VT + NP_2 --> NP_2 + Aux + be + part. + VT + by + NP_1
e.g. The mouse killed the fly --> The fly was killed by the mouse
Operations: tr, ad, ad, tr, ad
Units: 5
Note: It has been found necessary to distinguish passive constructions from so-called "impersonals." That is, sentences like
This is called a door.
This is stuck.
The cars are magnetized.
even though each contains "be + part + VT," are not scored as passives. Rather, it is assumed that the verbs are "to be called," "to be stuck,"

"to be magnetized." There are two general criteria for distinguishing passives from impersonals:

1. Passives optionally have a "by + NP" phrase attached; impersonals normally do not. i.e., we would not say "This is magnetized by the man."
2. A passive sentence corresponds to an "active" sentence with a definite specifiable subject, e.g.

The fly was killed (by me.)
I killed the fly.

Impersonals, when a corresponding "active" sentence is possible at all (which it often is not, e.g. "The door is stuck."), have indefinite, unspecifiable subjects, often expressed by "People" or an unspecified "they", e.g.

This is called a door.
People call this a door.
They call this a door.

Some other impersonals:

It's made for that track.
Everything was made of candy and cake.

A relatively common colloquial variant of T-passive substitutes "get" for "be", for example

The fly got killed.
He got hit by a car.

These are coded like normal passives, just as though "be" were present instead of "get".

18. T-passive-inf-a: NP_1 + Vto + to + VT + NP --> NP_2 + Vto + to + be + VT + part
(Note: Vto + to is often an Aux-s
e.g. The mouse has to kill the fly -->
The fly has to be killed.
Operations: tr, sub, ad, ad
Units: 4

19. T-passive-inf-b: NP_1 + VTo + to + VT + NP_2 --> NP_2 + Vto + to + be + VT + part + by + NP_1
e.g. The mouse has to kill the fly -->
The fly has to be killed by the mouse
Operations: tr, tr, ad, ad, ad
Units: 5

20. T + ?: X --> X + (question inflection)
e.g. Was John going --> Was John going?
or You know him --> You know him?
Operations: infl
Units: 1

21. T-Refl. X + Pronoun-m + Y --> X + Refl + Y (Pronoun -m: cf. T-obj)
e.g. He hits him --> He hits himself
You do it with you --> ... with yourselves
Units: 2
Operations: sub, ad
Note: in some cases + 1 must be scored for an optional element in the matrix as in
You do it with yourselves
I want to do it myself
but not in
He hits himself.

22. T-same as-a: NP + Aux + be + the + same --> NP + Aux + be + the + same (+ NP) + as + NP
e.g. It was the same --> It was the same as the other one
Operations: ad, ad
Units: 2

23. T-series-P: $X + Z + Y \rightarrow X + Z + Z^1 + Y$
where Z = NP, VP, N, V, adj, intens, adv, but not voc or interj.
e.g. a) Yesterday John went to the store -->
Yesterday John, Sam went .....
b) Yesterday John went to the store -->
Yesterday John went to the store, chopped the wood.
Operations: ad
Units: 1
Note: When scoring strings of adverbs, two adverb of the same type (e.g. adv-p + adv-p) are scored using T-series-p, but two adverbs of different types (e.g. adv-p + adv-m) are not: they are scored as additional optional units in the matrix. For example.
Now we could play with this first.
gets T-series-p (and also T-transp-adv) for two adv-t's (now, first), but
Now we could play with this over there.
gets + 2 in the matrix for an adv-t (now) and an adv-p (over there), and not T-series-p.

T-series-p adds only elements which maybe present in the matrix, and not elements added by

other T-rules. Therefore, when two interjections or vocatives are present in a single sentence, they are scored T-interj or T-voc twice, e.g.

Mommy, Mommy, (T-voc, T-voc) (2 units)
Hey, Mommy, hey. (T-voc, T-interj, T-voc) (3 units).

In a sentence such as

He doesn't laugh, doesn't cry, doesn't even talk,

T-series-p is applied twice to add "cry" and "talk" while T-neg and T-do are each applied three times, once for each occurrence of "doesn't."

In adding matrix elements, T-series-p may duplicate an element already present, as in

...very, very hot...

or it may add another element of the same type, as in

... a big, black dog...
... a man, a boy...

24. T-superl-a: $X + adj + Y \rightarrow X + \left\{\begin{matrix} adj + est \\ most + adj \end{matrix}\right\} + Y$

e.g. I am big --> I am biggest
I am beautiful --> I am most beautiful
Operations: ad
Units: 1
Note: + 1 for an optional element must be scored in the matrix for the adjective in

Show me the biggest one

but not in

I am biggest
I am most beautiful.

Note also that the remarks concerning "More" and "better" under T-compar-a also apply to "most" and "best" here.

25. T-superl-b: $X + adj + Y \rightarrow X + the + \left\{\begin{matrix} adj + est \\ most + adj \end{matrix}\right\} + Y$

e.g. I am big --> I am the biggest
I am beautiful --> I am the most beautiful
Operations: Ad, ad
Units: 2
(See notes for T-superl-a)

26. T-there: X + Nondef + Y + Aux + be + Adv-p -->
there + Aux + be + X + Nondef + Y + adv-p
e.g. Tomorrow a big cat will be in the sky --
There will be tomorrow a big cat in the sky
Operations: ad, tr
Units: 2
Note: Compare T-there with T-transp-adv in:
There they are
There he goes
T-there transposes the verb, while T-transp-adv does not. Therefore, in sentences beginning with "there," if the verb precedes the subject it will be coded T-there; if it follows, it is coded T-transp-adv. If the verb is missing it may be difficult to decide between T-there and T-transp-adv, e.g.
a) There another C.
b) There ball.
In such cases, when it is impossible to determine where the verb should be, the simplest derivation is assumed; and T-transp-adv, having a complexity of 1 unit, is scored, rather that T-there with the complexity of 2.

27. T-Transp-adverb: ...X + adv + Y ... --> { ... adv + X + Y ... / ...X + Y + adv ... }
where (...) indicates that other elements may be present e.g. I read books frequently --> I frequently read books, or Frequently I read books.
Operation: tr
Units: 1
Note: + 1 is usually scored in the matrix if an adverb is present, even though it has been transposed. The exception is the case where an adverb may follow "be" as a non-optional element, e.g. Here it is (from the matrix "It is here.) Certain adverbs (notably only and just) have more than one "natural" position within the sentence and are not given units for transposition when they are in one of these positions. Thus, none of the following sentences is scored T-transp-adv.

I see only letters.
I only see letters.
I see letters only.
I just have one.
I have just one.

In fact, note that final position, which is normal for most adverbs, is unnatural for "just", i.e. "I have one just" would not be considered grammatical Any adverb at the beginning of a sentence (except those accounted for by T-adv-sm, T-there, and T-invers) must be scored T-transp-adv. This holds for all of the following:

Only he could do that.
Now you do it.
Here you are.
After that we'll play with the house.
Now what's that name?

Sentences analogous to those produced by T-there, but with "here" in place of "there" must be scored T-transp-adv. i.e.

The book is here --> Here is the book

is scored T-transp-adv and T-transp-NP.
Occasionally sentences like the following are found:

Here's a piece here.
Here's one of the weights here.

These are scored T-series-P (to duplicate "here") and T-transp-adv.
(See also the notes for T-there.)

28. T-transp-Clause: If a clause can be assigned a "natural" position in the sentence, this transformation will be scored whenever the clause appears elsewhere in the sentence. This transformation scores 1 unit of complexity for the operation "tr".
e.g. I don't know why he did it --> why he did it, I don't know.

29. T-Transp-NP: ... X + NP + Y ... --> { ... NP + X + Y ... / ... X + Y + NP ... }
where (...) indicates that other elements may be present. e.g. I like to ride my bike --> My bike I like to ride.
Operations: tr
Units: 1

30. T-Transp-preart: X + Preart + NP + aux + V + Y --> X + NP + aux + V + preart + Y
where preart = all of
e.g. All of the children are going to the movies --> The children are all going to the movies
Operations: tr

Units: 1
Note: It should be remembered that a pre-article requires + 1 in the matrix

31. T-voc:

X + Y --> X + NP voc + Y
where NP voc is a vocative element such as a name
e.g. The tea is ready --> John, the tea is ready,
or The tea is ready, John
Operations: ad
Units: 1
Note: Vocatives may appear anywhere in the sentence and are never scored for transposition. Vocatives may be the names of inanimate objects.
e.g. C mon, car.
or may consist of more than one word, e.g.
Mr. Smith!
Get up, you lazy bum.
Vocatives never get T-series-p, and if several vocatives appear in a series, each one is scored T-voc. However, they may be scored T-conj-P, in sentences like
OK, Lisa and Eric, come on.

32. T-VT:

X + VT + prt or Comp + NP --> X + VT + NP + Prt or Comp
e.g. I put away them --> I put them away
They considered wrong the U.S. Senator --> They considered the U.S. Senator wrong (preceded by T-VT3)
Operations: tr
Units: 1
Note: It is often difficult to distinguish particles (prts) from prepositions (preps.) A particle is part of the verb; a preposition is not
Consider the following:

1a. Take that off.
1b. Take that off him.
2a. Get it on.
2b. Get on it.
3. The horse was stepping on him.
4. The horse pushed it down.
5. Turn it around.
6. Pick it up.
7. Push it all the way in.
8. Give it back to me.

One criterion is that a prt. may be transposed to a position following the object (as in examples 1a, 2a, 4, 5, 6, 7, 8) whereas a prep. cannot. (e.g. 2b, where the meaning is changed if "on" is transposed). Another is that a prep can have an object independent of the object of the verb. (This is usually obvious only when there are two objects present, as in example 1b) However, the decision must sometimes be left to the intuition of the coder in the absence of unequivocal criteria e.g. example 3, may be viewed either as V (step on) + object (him) or as V (step) + adv-p (on him). Our tendency has been to assume prt., rather than prep, in such cases.

33. T-wh-adv-m: X + adv-m + Y --> how + X + Y
e.g. He ran quickly --> how he ran (may be followed by T-yes/no, T-do, and T-?)
Operations: ad, sub
Units: 2
Note: + 1 must be scored in the matrix for adv-m.

34. T-wh-adv-no: a) X + number + N + Y --> how many + N + X + Y
b) X + Number + Y --> how many + X + Y
e.g. a) Two men robbed the bank -->
b) How many men robbed the bank
Two robbed the bank --> How many robbed the bank.
(Normally followed by T - ?, often also by T-yes/no.)
Operations: ad, sub.
Units: 2
Note: This transformation is also used to score sentences with "how much" instead of "how many", e.g. How much is that?
If "how much/how many" precedes a noun (e.g. "how many men," "how much money"), it is scored in the matrix; otherwise, it is not (e.g. "How much is that?" "How many are there?")

35. T-wh-adv-p: X + adv-p + Y --> where + X + Y
e.g. I was in the park today --> where I was today. (T-yes/no and T-? can follow)
Operations: ad, sub
Units: 2
Note: T-wh-adv-p normally requires + 1 in the matrix, e.g. "Where does the train go?"
"Where do you work?"
except where the underlying kernel is of the

type "NP + be + adv-p,"
e.g. where were you? (from the matrix "you were in the park.")

36. T-wh-adv-t: X + adv-t + Y --> when + X + Y
e.g. I was in the park today --> When I was in the park (T-yes/no and T-? can follow)
Operations: ad, sub
Units: 2
Note: An adv-t usually must be scored + 1 in the matrix, except in sentences of the type "NP + be + adv," e.g. "It was yesterday" (--> "when was it?" by T-wh-adv-t, T-yes/no, & T-?).

37. T-wh-art: X + art + N + Y --> which or what + N + X + Y
e.g. Slowly the man walks --> which man slowly walks. (T-yes/no & T-? may follow)
Operations: ad, sub, tr
Units: 3
Note: "which" can only be derived by T-wh-art, even if no noun is present, e.g.
Which is that?
must be scored T-wh-art. However, "what" can be the result of either T-wh-art or T-wh-NP. That is,
What book is that?
is a T-wh-art, but
What is that?
is a T-wh-Np. Occasionally ambiguity results, so that
What quacks?
may be derived from
The duck quacks
by T-wh-NP, or from
(Did you hear) those quacks?
by T-wh-art. The underlying sentence can usually be determined from the context, but if not, the simplest deviation is assumed, i.e. T-wh-NP (2 units) instead of T-wh-art (3 units).
Some further examples of T-wh-art:
What color is that?
What time is it?
What toys?

38. T-wh-intensifier: X + int + $\left\{\begin{matrix}adj\\adv\end{matrix}\right\}$ + Y -->

How + $\left\{\begin{matrix}adj\\adv\end{matrix}\right\}$ + X + Y

e.g. She is very pretty -->
How pretty she is
Operations: sub, ad, tr
Units: 3
Note: T-wh-intensifier requires + 1 in the matrix for an intensifier.

39. T-wh-NP: X + NP + Y --> WHO or WHAT + X + Y
("who" if NP is animate," "what " otherwise)
e.g. Someone stole $17 --> who stole $17
(T-yes/no & t - ? can follow)
Operations: ad, sub
Units: 3
Note: cf: "what" as derived by T-wh-art. Some other examples of T-wh-NP:

a. You think - NP - is in the house (i.e. preceded by T-sub-a) --> What do you think is in the house?
b. Something happened --> What happened?
c. That's the book --> That's the what?
d. He lost it --> Who lost it?
e. This part of the train is called the engine --> What is this part of the train called?
f. Something is the matter --> What's the matter?
g. He's doing something --> What's he doing?
j. You've got something --> What have you got?
i. You want the book --> You want what? or You want the what? or What do you want?

40. T-wh-possessive: X + poss + N + Y --> whose + N + X + Y
where poss. is either a possessive pronoun or a possessive noun derived by means of T-pos.
e.g. This is his hat --> whose hat this
the man's

is.
(T-yes/no & T -? can follow).
Operations: sub, ad, tr
Units: 3

41. T-why: X + adv-purp. + Y --> why + X + Y
where adv-purp is an adverbial of purpose,
e.g. "for that reason"
e.g. He did it for that reason --> why he did it
(T-yes/no & T-? may follow)

Note: T-why normally requires + 1 in the matrix for an optional adverb. A colloquial variant of T-why substitutes "how come" for "why". There are differences between "why" and "how come" in that the former normally requires T-yes/no and T-do, while the latter does not. Compare the following:

a. Why does it balance?
   How come it balances?
b. Why does something happen?
   How come something happens?

However, this peculiarity concerns the accompanying transformations (T-yes/no and T-do) and not the operation of the T-why rule itself, so "how come" can be scored like "why" by using T-why.

42. T-word: This transformation scores one unit of complexity when a word performs a function that is not normal for that type of word. e.g. an adjective functioning as subject ("Pretty" is a pretty word) or an animate noun as an inanimate ("John" is a noun), etc.
T-word is most common with verbs like "say," "tell," etc. for example: Say bye-bye
This says Matthew
I said no
But the following are not scored T-word:
That says your name
That's a U
That's a five.
Notice that T-sep is similar, except that it is a double-base transformation that inserts entire clauses, in the form of quotations, after "say," "tell", etc.

43. yes/no:

a) NP + tense $\left\{\begin{matrix}\text{have}\\ \text{M}\\ \text{be}\end{matrix}\right\}$ + X -->

tense = $\left\{\begin{matrix}\text{have}\\ \text{M}\\ \text{be}\end{matrix}\right\}$ + NP + X

b) NP + tense + verbal --> tense + NP + verbal
e.g. a) John was going there --> was John going there
b) John went to school --> (past) + John go to school (requires T-do).

Operations: tr
Units: 1
Note: T-yes/no presents some rather special problems when the present tense ending is missing and with certain wh-questions. In the sentence:

That go there?

there are two possible explanations for the absence of the ending on "go" depending on the assumption made about the kernel: either

1. The ending was not present in the underlying kernel, or
2. It was present in the kernel and was eliminated by one of the T-rules. If we accept assumption (1), the derivation would be:

K: that go there
T-?. That go there?

However, assumption (2) would support the following derivation:

K: that - pres + go -there.(pres + go = goes).
T-yes/no: pres-that-go-there.
T-?: pres-that-go-there?

Then, if T-do is applied, we get: do+ pres-that-go-there? --> Does that go there?
However, if T-do is not applied, the tense element "pres" is lost and "That go there?" is all that remains. No objective criteria have been found for deciding between these two alternatives and it was decided to assume the simplest derivation - in t his case, the one in which "pres" was absent in the underlying kernel and T-yes/no was not applied. Thus the sentence "That go there?" receives only 2 + 1 units in the matrix and 1 unit for T-?.
The principle of assuming the simplest possible derivation also applies to wh-questions such as

Who drives you?
What came off?
Whose name is this?
What happened?

Where the sentence can be formed either with or without T-yes/no, e.g.

K; He-pres-drive-you
T-yes/no: pres-he-drive-you
T-wh-NP: who-pres-drive-you
T-?: who-pres-drive-you?
--> who drives you?

where the effect of T-yes/no is nullified by the application of T-wh-NP. Although the grammar as a whole will be simplified if we can say "T-yes/no must always precede T-wh-NP", instead of trying to define those circumstances when T-yes/no does apply and when it doesn't, it was decided not to score T-yes/no unless the sentence could not be derived without it (as in "where is he?" where "is" has been transposed to a position preceding the subject). Similarly, in the sentence

What's this?

We can assume either of the following kernels:

X is this
this is X

} where X becomes as "what" by T-wh-NP.

Since the former will require a simpler derivation to get "what's this?" it is the assumed underlying kernel.

One further question related to T-yes/no has to do with "some" and "any". Consider the following sentences:

Is anybody at home?
Do you have any?
I don't have any.

One's first inclination would be to assume the following underlying kernels:

Somebody is at home.
You have some.
I have some.

However, a no satisfactory "some --> any" rule has been written, and some linguists have suggested that such a rule is impossible. It has therefore been necessary to assume the following kernels for the above examples:

Anybody is at home.
You have any.
I have any.

44. T-yes/no 2A: NP + tense + {M, have, be, verbal} + X -->

NP + tense + {M, have, be, verbal} + X + tense

+ {M, have, be, Ø} + not + pronoun

where Ø signifies a zero element and indicates that T-do must be applied
e.g. Children are devils --> children are devils, aren't they
He eats a lot --> He eats a lot, doesn't he
Operations: ad, ad, ad
Units: 3

45. T-yes/no 2B: NP + tense + {M / have / be / verbal} + not + X -->

NP + tense + {M / have / be / verbal} + not + x +

tense {M / have / be / Ø} + pronoun

(for meaning of Ø, see T-yes/no 2A)
e.g. Children don't speak well -->
Children don't speak well, do they
Operations: ad, ad
Units: 2

C. DOUBLE BASE TRANSFORMATIONS *

In this list, the units indicated for each transformation include the transformational operations and the insert kernel. Units for the matrix and any optional units in the insert kernel are scored separately. As with Single-Base Transformations, the number of units indicated here may be adjusted if the T is not applied exactly according to the rule.

---

* For a more up-to-date treatment of double base transformations, see N. Chomsky, Aspects of the Theory of Syntax, Cambridge, M.I.T. Press, 1965; and R. Jacobs and P. Rosenbaum, English Transformational Grammar, Waltham, Blaisdell, 1968.

The symbols W, X, Y, Z may again stand for any or no element unless otherwise specified, and ad, sub, tr, infl indicate addition, subtraction, transposition, and inflection. "V" includes "be" and "verbal", and "comp" is a dummy symbol indicating that it ir to be replaced by an insert kernel.

46. T-as:

matrix: NP_1 + aux + V + $\begin{Bmatrix} adj \\ adv \end{Bmatrix}$

insert: NP_2 + aux + V + $\begin{Bmatrix} adj \\ adv \end{Bmatrix}$

where V is a verbal or "be".

result: NP_1 + aux + V + as + $\begin{Bmatrix} adj \\ adv \end{Bmatrix}$ as + NP_2

+ aux + V

e.g. a) matrix: he is old
insert: I am old
result: He is as old as I am

Operations: K, ad, ad, sub
Units: 5
Note: The sentences
a) Let's get it back together just as it was
b) He is as old as you.
each get 4 units for T-as instead of the usual 5: example a) because the adj/adv is missing, b) because the second verb is missing. Compare
I'm working as a case worker,
where "as" belongs to an adverbial phrase (i.e. is scored + 1 in the matrix), not to T-as.

47. T-apos:

matrix: X + NP + Y
insert: NP + aux + be + Z
result: $\begin{Bmatrix} X + NP + Z + Y \\ X + Z + NP + Y \end{Bmatrix}$

e.g. matrix: I see my sister every day.
insert: My sister is Jane.
result: I see my sister Jane very day. or
I see Jane, my sister, every day.

Operations: K', ad
Units: 3
Note: Apositions have several possible locations within the sentence and will not be scored T-transp.

48. T-conj-K:

matrix: NP + VP
insert: NP' + VP'
result: NP + VP + conj + NP' + VP'
where conj = and, or, but, etc.
e.g. matrix: John goes to school.
insert: his brother stays home.
result: John goes to school, and (or, but, etc.) his brother stays home.
Operations: K', ad
Units: 3
Note: Like the other series and conjunction transformations, T-conj-K does not add or duplicate transformational elements; it can only add elements which may be present in a kernel sentence. Therefore in

I didn't eat and I didn't sleep,

T-do and T-neg must each be scored twice. The same principle holds true for conjoined sentences containing imperatives, and each of the following sentences is scored T-imper-a twice, and T-conj-K (not T-conj-P):

Look around the room and find something else.
Come and tell me what this is.
Pick up your toys and put them away.

49. T-compar-b:

$$\text{matrix: } NP_1 + \text{aux} + V + \left\{\begin{matrix}\text{adj}\\ \text{adv}\end{matrix}\right\}$$

$$\text{insert: } NP_2 + \text{aux} + V + \left\{\begin{matrix}\text{adj}\\ \text{adv}\end{matrix}\right\}$$

$$\text{result: } NP_1 + \text{aux} + V + \left\{\left\{\begin{matrix}\text{adj}\\ \text{adv}\end{matrix}\right\} + \text{er} \quad \text{more} + \left\{\begin{matrix}\text{adj}\\ \text{adv}\end{matrix}\right\}\right\} + \text{than} + NP_1 + \text{aux} + V$$

e.g. matrix: I am big
insert: He is big
result: I am bigger than he is
Operations: K', ad, ad, sub
Units: 5
Note: Sentences like the following, with the second verb missing, are fairly common:

He is bigger than the horse.
I eat more than you.

These are given 4 units for T-compar-b, instead of the usual 5.

50. <u>T-indir-q:</u> matrix: X + NP + Y

insert: { how, when, where, who, what, which, why } + N + Z (i.e. preceded by T-wh-adv m, wh-advt, wh-adv P, wh-NP, wh-art or why, but <u>not</u> T-yes/no)

result: X + { how, when, where, who, what, which, why } + Z + Y

e.g. matrix: I saw something yesterday
insert: What they were building (from "They were building something" by T-wh-NP)
Result: I saw what they were building yesterday

Operations: <u>K'</u>, <u>sub</u>, <u>ad</u>
Units: <u>4</u>
Note: Indirect questions containing "how", "when", "where", or "why" often require + 1 in the insert kernel for an adverb, e.g.

I don't know where you found it.

contains the insert kernel

You found it + adv-P

which contains an optional adv-P. In such cases + 1 is scored in the insert kernel only, not in the matrix. The following are scored T-indir-q.

Look what happened.
Look what I did.
That's what you found over there.
Do you know what color that is?
(Insert = "that is that color" followed by T-wh-art.)
Don't tell him what I did.
He never changes what he is.

We have to find out what and four makes ten so we can make it balance.
(matrix: We have to find out + NP)
(insert: Six and four makes ten.)

51. T-indir.q-b: Matrix: X + NP + Y

Insert: { what, when, where, how, who, which } + N + Z (i.e. preceded by T-wh-NP, wh-advt, wh-advp, wh-advm, wh-art, but not T-yes/no.)

Result: X + { what, when, where, how, who, which } to + Z + Y

e.g. Matrix: I saw something
Insert: What I build
Result: I saw what to build
Operations: K', sub, sub, ad
Units: 5
The notes to T-indir q also apply to T-indir q-b.

52. T-ing-a:
matrix: X - det + N - Y
insert: NP - Aux - Z
result: X - Det + N + ing = Z - Y
e.g. matrix: We noticed the N yesterday
insert: Mary shouted
result: We noticed the shouting yesterday
Operations: K', sub, ad, ad
Units: 5

53. T-ing-b:
matrix: X - NP - Y
insert: NP - Aux - Z
result: X - NP - Pos - ing - Z - Y
e.g. matrix: We noticed - NP - yesterday
insert: Mary shouted at the top of her lungs
result: We noticed Mary's shouting at the top of her lungs yesterday.
Operations: K', sub, ad, ad, ad, ad
Units: 7

54. T-ing-c:

matrix: X + NP + Y
insert: NP + aux + Z
result: X + NP + ing + Z + Y
e.g. matrix: We noticed - NP - yesterday
insert: Mary shouted at the top of her lungs
result: We noticed Mary shouting at the top of her lungs yesterday
(N.B. This is the same as T-ing-b but without pos in the result.)
Operations: K', sub, ad, ad, ad
Units: 6
The following is also a T-ing-c:
You have the engine pulling it.

55. T-ing-adj:

matrix: NP + VP
insert: NP + aux + V
result: NP + VP + V + ing
e.g. matrix: I go up; the stairs
insert: I laugh
result: I go up the stairs laughing
Operations: K', a, ad, sub
Units: 4

56. T-ing-inf:

matrix: X + Vto + TO + V + Y
insert: X + V + Z
result: X + Vto + Vt ing + Z + Y
where Vto may be an Aux-s or the result of T-to-NP - b
e.g. matrix: He started to + V
insert: He laughed
result: He started laughing
Operations: K', sub, ad
Units: 4

57. T-parenth:

matrix: X + Y
insert: NP' + VP'
result: X + (parenthinflection) + NP' + VP' + Y
e.g. matrix: Yesterday I was very tired
insert: I had worked 18 hours.
result: Yesterday (I had worked 18 hours) I was very tired.
Operations: K', infl, ad
Units: 4
Note: Parenthetical expressions can appear almost anywhere in the sentence and will not be scored T-transp.

58. T-pos:

matrix: X + Det + N + Y
insert: NP + aux + have + Det + N
result: X + NP + Pos + N + Y

where pos ≠ possessive pronoun

e.g. matrix: I want to have the book now.
insert: The general has a book
result: I want to have the general's book now.

Operations: K', sub, ad, ad
Units: 5

59. T-rel-a:

matrix: X + NP + Y
insert: Z + NP + W
result: X + NP + who, which or that + Z + W + Y

e.g. matrix: Yesterday we noticed the birds.
insert: On Friday the birds escaped from the zoo.
result: Yesterday we noticed the birds who on Friday escaped from the zoo.

Operations: K', ad, ad
Units: 4

60. T-rel-b:

matrix: X + V ... + NP
insert: Y + V + NP
result: X + V ... + NP + Y

e.g. matrix: Show me the picture
insert: You drew the picture
result: Show me the picture you drew.

Operations: K', sub
Units: 3
The following are also T-rel-b's:
Do it any way you want
This is the way it goes.

61. T-rel-when:

matrix: W + NP_1 + X
insert: Y + adv-t + Z
result: W + NP_1 + when + Y + Z + X or W + NP_1 + X + when + Y + Z
where adv-t = prep + NP_1

e.g. I remember the day
The war ended on the day
I remember the day when the war ended

Operations: K', ad, ad
Units: 4
Note: T-rel-when normally requires + 1 in the insert kernel for an adv-t.

62. T-rel-where:

matrix: W + NP_1 + X
insert: Y + adv-p + Z
result: W + NP_1 + where + Y + Z + X or W + NP_1 + X + where + Y + Z
where adv-p = prep + NP_1

e.g. matrix: I found a nest in my yard
insert: Birds hatch eggs in a nest in the spring
result: I found a nest where birds hatch eggs in the spring in my yard or I found a nest in my yard where birds hatch eggs in the spring.
Operations: K', ad, ad
Units: 4
Note: T-rel-where may require + 1 in the insert kernel.

63. T-same as -b: matrix: NP + aux + be + the + same (+ NP)
insert: NP + V + comp
result: NP + aux + be + the + same (+ NP) + as + NP + VP
e.g. matrix: It was the same
insert: I thought + comp
result: It was the same as I thought
Operation: K', ad, sub
Units: 4

64. T-same as-c: matrix: NP + aux + be + the + same (+NP)
insert: NP + V + Comp
result: NP + aux + be + the + same (+NP) + as + when + NP + VP
e.g. matrix: It was the same
insert: I had thought + comp
result: It was the same as when I had thought
Operations: K', ad, ad, sub
Units: 5

65. sep: matrix: X + aux + vq + NP + Y
insert: Z
result: X + aux + Vq + (quotation inflection) + Z or (quot. infl.) Z + X + aux + Vq + Y
where Vq = say, tell, state, etc.
e.g. matrix: They said + NP + this morning
insert: What an ugly sunrise!
result: They said this morning, "what an ugly sunrise!" Or "What an ugly sunrise!" they said this morning.
Operations: K', sub, ad, ad
Units: 5

66. series-K:

matrix: - NP + VP
insert: -NP' + VP'
result: - NP + VP + NP' + VP'
e.g. matrix: - John goes to school
insert: - His brother stays home.
result: - John goes to school, his brother stays home.
Operations: K'
Units: 2
Note: T-series-p adds only elements which may be found in matrix sentences and any transformational elements or operations found in both kernels must be scored twice for that transformation (cf. T-conj-K). The most common example of this is in imperatives like

Pick it up, pick it up

where T-imper-a is scored twice, in addition to T-series-p. Note also that in this example T-Vt must also be scored twice. Some other T-series-K's:

See that train, it's moving.
Say, this is a magnet letter.
Hush, hush now.

67. T-sub-a:

matrix: X + NP + Y
insert: NP' + VP
result: X + NP' + VP + Y
e.g. matrix: He knew + NP + in his heart
insert: They would win
result: He knew they would win in his heart
Operations: K', sub, ad
Units: 4

68. T-sub-b:

matrix: X + NP + Y
insert: NP' + VP
result: X + sub_1 + NP' + VP + Y
where sub_1 = that, if, because, etc.
e.g. matrix: He knew + NP + in his heart
insert: They would win
result: He knew that they would win in his heart.
Operations: K', sub, ad, ad
Units: 5
The following are also scored T-sub-b:

See if the roof comes off.
See if you can find it.
That is because it's small.
Let's see if you're a big boy.

69. T-sub-c:

matrix: NP + aux + be + comp
insert: S
result: NP + aux + be + that + S
where S = a sentence
e.g. matrix: The story is + comp
insert: He killed her
result: The story is that he killed her.
Units: 4

70. T-sub-d:

matrix: NP + { aux + be + adj / aux + V + adv }

insert: NP' + VP'

result: NP + { aux + be + adj / aux + V + adv } + that + NP' + VP'

e.g. matrix: It was so good
insert: I couldn't stop
result: It was so good that I couldn't stop
Units: 4

71. T-sub$_2$:

matrix: NP + VP
insert: NP' + VP'
result: Sub$_2$ + NP' + VP' + NP + VP
where Sub$_2$ = when, if, whenever, so, since, because, so that, etc.
e.g. matrix: I will go there
insert: The babysitter comes.
result: Whenever the babysitter comes, I will go there.
Operations: K', ad, ad
Units: 4
Note: The Sub$_2$ clause is not restricted to the position at the beginning of the sentence before the NP of the Matrix. There are several points within the sentence where it may occur, none of which will be considered preferable. Thus Sub$_2$ will not come under any transposition transformation The following is also scored T-sub$_2$:
The car is going to move so that the train can go on the track.

72. T-subj a:
73. T-subj b :
74. T-subj c:

Omit

75. T-to-a:

matrix: it + aux + be + adj + comp
insert: NP + aux + X
result: it + aux + be + adj + to + X
e.g. matrix: It will be unnecessary - comp
insert: They will chop wood.
result: It will be unnecessary to chop wood
Operations: K', sub, ad, ad
Units: 5

76. T-to-b:

matrix: it + aux + be + adj + comp
insert: NP + aux + V
result: it + aux + be + adj + for + NP + to + X
e.g. matrix: It will be necessary - comp
insert: They will chop wood
result: It will be necessary for them to chop wood.
Operations: K', sub, ad, ad, ad, ad
Units: 7

77. T-to-adj-a:

matrix: X + NP_1 + Y
insert: NP_2 + V + NP_1
result: X + NP_1 + to + V + Y
e.g. matrix: He makes hats
insert: Tney look at hats
result: He makes hats to look at
Operations: K, sub, sub, ad
Units: 5

78. T-to-adj-b:

matrix: X + NP_1 = Y
insert: NP_2 + V + NP_1
result: X + NP_1 + Y + for + NP_2 + to + V + Y
e.g. matrix: He makes hats
insert: They look at hats
result: He makes hats for them to look at
Operations: K', ad, ad, ad
Units: 5
Another example of T-to-adj-b:
matrix: There is some hay in the barn (from "Some hay is in the barn" by T-there)
insert: It eats some hay
result: There is some hay in the barn for it to eat.

79. T-to-adv:

matrix: NP + VP
insert::NP' + V + X
result: NP' + VP + TO + V + X
e.g. matrix: It is good
insert: I am here
result: It is good to be here
Operations: K', ad, ad
Units: 4
Note: T-to-adv has been applied twice to the following sentence:
It is time to come to eat dinner.
The original matrix and insert kernels are
It is time
NP' - comes.
which, on the first application of T-to-adv gives
It is time to come.
This then serves as the matrix for the second insert kernel
NP' - eats dinner
which gives
It is time to come to eat dinner
after the second application of T-to-adv.

80. T-to-NP-a:

matrix: Z + comp + VP
where subject of comp is not 1 unit or X
insert: NP + V + X
result: Z + TO + V + X + VP
e.g. matrix: always, comp, is heroic
insert: men die for the country
result: always, to die for the country is heroic
Operations: K', ad, ad, tr
Units: 5

81. T-to-NP-b:

matrix: NP + VT + comp
insert: NP + V + X
result: NP + VT + TO + V + X
e.g. matrix: he continues + comp
insert: he + aux + go + to school
result: he continues to go to school
Operations: K', ad, ad
Units: 4
Note: the similarity between T-to-NP-b and the aux-s (see above, section II B). The following are all aux-s's:
try to
want to
have to
be going to

Compare the foregoing with
The former is getting ready to
bang the table
which is a T-to-NP-b.

82. T-to-NP-c: matrix: NP + aux + be + comp
insert: NP' + V + X
result: NP + aux + be + to + V + X
e.g. matrix: the object is + comp
insert: He hit the ball
result: The object is to hit the ball
Operations: K', ad, ad
Units: 4

83. To-Purpose-a: matrix: NP + VP_1
insert: NP + VP_2
result: NP + VP_1 + TO + VP_2
e.g. matrix: He makes hats
insert: He earns money
result: He makes hats to earn money
N.B. can be VI or VT, e.g. Stays in N.Y. to earn ---
Operations: K', sub, ad, ad
Units: 5
Other examples:
The cow goes in to be milked.
You have one minute to complete the job.

84. T-to-purpose-b: matrix: NP + VP_1
insert: NP + VP_2
result: NP + VP_1 + in order + to + VP_2
e.g. matrix: He makes hats
insert: He earns money
result: He makes hats in order to earn money
Operations: K', sub, ad, ad, ad
Units: 6

85. T-VTO-a: matrix: NP + aux + VTO + comp + NP'
where VTO = let, make...
insert: NP' + aux + X
result: NP + aux + VTO + X + NP'
e.g. matrix: They let + comp + the girls
insert: the girls + past + dance
result: They let dance the girls . (requires T-VT)
Operations: K', sub, ad
Units: 4

Other examples (T-Vt has already been applied):
You let it go away .
I'm watching you do this puzzle.

86. T-VTO-b:

matrix: NP + aux + VTO + comp + NP'
insert: NP' + aux + X
result: NP + aux + VTO + to + X + NP'
where VTO = allow, want...
e.g. matrix: They allowed + comp + the girls
insert: The girls danced
result: They allowed to dance the girls
(requires T-VT)
Operations: K', sub, ad, ad
Units: 5
Other examples (T-VT has already been applied):
Cows like people to ride on them .
You want Mommy to do it?
I want everybody to play with it.
You want her to get up?
I've gotta get him to play with that puzzle.
I want you to stop that.

87. T-adv VT_3:

matrix: NP + aux + VT_3 + comp + NP'
(VT_3 --> think, consider...)
insert: NP' + aux + be + substantive
result: NP + aux + Vt_3 + substantive + NP'
(substantive = an NP or an adj)
e.g. matrix: He thinks + comp + the elephants
insert: The elephants are intelligent
result: He thinks intelligent the elephants
(requires T-VT)
Operations: K', sub, ad
Units: 4

88. T-want-a:

matrix: Y + NP_1 + V want + NP_2 + comp
(V want --> want, have, get, be, find, order, need, etc.)
insert: NP_2 + tense + be + VT + ed + X
(i.e. may be preceded by T-passive)
result: Y + NP_1 + V want + NP_2 + VT + ed + X
e.g. matrix: I want the man + comp
insert: The man is searched
result: I want the man searched
Operations: K', sub, ad
Units: 4
The following are also scored T-want-a:

a) They found a little house made out of candy
   [matrix: They found a little house + comp.
   insert: A little house is made out of candy]
b) This is something made out of wood.
c) I have a book called War and Peace.
   [matrix: I have a book + comp
   insert: A book is called War and Peace.]

89. T-want-b:

matrix: Y + NP_1 + V + want + NP_2 + comp
where V want = want, have, get, order, need, etc.
insert: NP_2 + tense + be + VT + ed + X
(i.e. may be preceded by T-passive)
result: Y + NP_1 + V want + NP_2 + to + be + VT + ed + X
e.g. matrix: I want the man + comp
insert: The man is searched
result: I want the man to be searched
Operations: K', sub, ad, ad, ad
Units: 6

IV. SCORING CONVENTIONS

When dealing with natural conversational speech, especially that of preschool children, one encounters a great number of utterances which do not quite fit a grammar of Modern Standard English. It has therefore been necessary to make some arbitrary decisions about how to score such utterances so that some degree of reliability can be obtained. A list of the most common of these conventions follows (the number of units to be scored in the matrix (m) and the T-rule to be used in scoring are indicated whenever appropriate):

SENTENCE	SCORING
Adv-P --> over there, right over here...	
All gone.	M = 1 unit
All right.	M = 1 unit
be able to = aux-s	
Because + ...	M =0 , T-sub 2
Beep beep. choo choo. Toot toot.	not scored
belong + adv-p	not + 1
Better: You better do it.	+ 1 unit for adv, (Not T-compar)
Can: I can.	M = 1 + 1 units
I can't.	M = 1 + 1, T-neg
Fire engine	M = 1 (not + 1 for adj).

SENTENCE	SCORING
Get + adv-P = be + adv p (not +1)	
Get it mixed up.	T-VT_3
Here doll.	M = 1 + 1, T-transp adv
Hmm? Huh?	M = ∅, T-interj, T-?

Imper a + Imper a (not Imper b or c) - use T-conj K, not T-conj p.
Interjection --> thank you, hello, please, yes, no, good night, welcome
Invers --> and, or, but...
Just: NP + just + VP not scored transp adv

Keep (e.g. Keep doing)≈ continue	T-ing inf
"let" in imper is always imper b or c	
Like: That's like a fire engine	V = to be like
That looks like a fire engine.	V = to look like
Like to. = aux-s	
Mommy.	T-voc
Mommy?	T-voc - T-?
Never mind = interj	
No more.	M = 2 + 2
Now...	T-transp adv

{NP + VT; NP + be} w/o {obj; subst; adv-p} - M = 2

NP + {obj; subst; adv} w/o V or be - M = 1

One.	M = 1
Put: NP + put + obj + adv-p	M = 2 + 1
Prt: sit down, lie down, take out, play with, etc. (see note on T-VT, p.32)	
Right:..., right?	interj
# Right? #	M = 1 ? - 1
Say: say + 1 word	T-word
say + clause	T-sep
See:	M = 1
See?	M = 1 T-? - 1
So...	sub 2

T-series does not apply to interj or voc

Turn it upside down.	T-VT_3
Sit down	V + prt, not + 1
Sit + adv-p (e.g. sit here, sit on the chair)	1 + 1
Some = article (not scored + 1 for adj)	
Something else	1 + 1

SENTENCE	SCORING
There: there + v + NP there + NP + V	T-there T-transp adv
Want to = aux-s	

SENTENCE	SCORING	UNITS
Wh-questions:		
How?	M = 1 T-wh adv m -2 T-? - 1	4 units
What?	M = 1 T-wh NP -2 T-? - 1	4 units
What else?	M = 1 + 1 T-wh NP -2 T-? - 1	5 units
What's that?	M = 2 T-wh NP -2 T-? - 1 (assumes M= NP + is that)	5 units
What's this called?	V = to be called M = 2 T-yes/no - 1 T-wh NP - 1 T - ? - 1	5 units
Where?	M = + 1 T-wh adv p-2 T-? - 1	4 units
Where is it?	M = 2 T-yes/no 1 T-wh-adv-p 2 T-? - 1	6 units
Who?	M = 1 wh NP -2 T-? - 1	4 units
Why?	M = 1 why -2 T- ? - 1	4 units

1. Do not score T-yes/no if the sentence could be formed without it. In sentences of the type wh + be + NP or wh + be + adv/adj, T-yes/no is not scored (e.g. What's that? Who's this? Who's there?)
2. If "be" or verbal is absent do not score yes/no (e.g. Where man?)
3. If "do" and a verbal are present, score yes/no (e.g. Where does he go?)

V. SCORING SENTENCE COMPLEXITY

A. GENERAL

In order to illustrate the process we are attempting to quantify, it may be useful at this point to illustrate the step-by-step procedure by which the rules presented in the foregoing sections produce ("generate") sentences.[5] The following sentences will serve as illustrations:

a. Don't tell him what I did.
b. Why don't you look around and come back and work on it some more.
c. He'll listen to an explanation and he will sometimes repeat the explanation, and it doesn't satisfy him and he will continue asking "why."

The lexical rules, which substitute words for class symbols, and the phonological rules, which give the sentences their final phonetic form, are irrelevant here and have been omitted.

---

5 This does not mean that the speaker goes through these steps to produce sentences; we are dealing with the linguistic, rather than the psychological, complexity of utterances. It is assumed, however, that the linguistic complexity we are attempting to define is in some way analogous to a psychological process (or processes) which we are unable to measure directly.

Rule	Result	Units
(a)		
S --> NP + VP	NP + VP	2
VP --> Aux + verbal	NP + aux + verbal	
aux --> tense	NP + tense + verbal	
verbal --> VT + NP	NP + tense + VT + NP	
NP --> Pron.	Pron. + tense + VT + NP	
Pron --> Pers	Pers + tense + VT + NP	
	You + tense + tell + NP	
T-indir obj	You + tense + tell he + NP	1
T-obj	You + tense + tell him + NP	1
T-neg	You + tense + not tell him + NP	1
T-do	You don't tell him + NP	1
T-imper-a	Don't tell him + NP	2
(Formation of insert)	I did + NP	(Units included in T-indir-q)
T-wh-NP	What I did	2
T-indir q	Don't tell him what I did	4
	Total units:	14
(b)		
S --> NP + VP	NP + VP	2
VP --> aux + verbal + adv	NP + aux + verbal + adv	+ 1 (for adv)
Adv --> adv purp	NP + aux + verbal + adv-purp	
Aux --> tense	NP + tense + verbal+adv-purp	
Verbal --> VI	NP + tense + VI + adv-purp	
NP --> Pron	Pron + tense + VI + adv-purp	
Pron --> Pers	Pers + tense + VI + adv-purp	
	You + tense + look around + for that reason	
T-why	why you + tense + look around	2
T-yes/no	why + tense + you look around	1
T-do	why + do + tense + you look around	1
T-neg	why don't you look around	1
T-conj-p	why don't you look around and come back	2
T-conj-p	why don't you look around and come back and work on it some more	2 + 1
	total units	13

Rule	Result	Units
(c)		
S --> NP + VP	NP + VP	2
VP --> AUX + verbal	NP + aux + verbal	
aux --> tense + M	NP + tense + M + verbal	+ 1 (for M)
verbal + VT + NP	NP + tense + M + VT + NP	
NP --> Pron	Pron + tense + M + VT + NP	
pron --> PERS	Pers + tense + M + VT + NP	
NP --> Det + N	Pers + tense + M + VT + Det + N	
Det --> art	Pers + tense + M + VT + art + N	
art --> nondef	Pers + tense + M + VT + nondef + N	
	He + tense + will + listen to + an + explanation	
T-conj K	He will listen to an explanation and he will repeat the explanation sometimes	3 + 2 (+1 for "will" + 1 for "sometimes")
T-transp-adv	He will listen to an explanation and he will sometimes repeat the explanation	1
T-conj K	He will listen to an explanation and he will sometimes repeat the explanation and it satisfies he	3
T-obj	... and it satisfies him	1
T-neg	... and it + tense + neg + satisfy him	1
T-do	... and it doesn't satisfy him	1
T-conj K	He will listen to an explanation and he will sometimes repeat the explanation and it doesn't satisfy him and he will continue	3 + 1 (+1 for "will")

Rule	Result	Unit
T-ing-inf	... and he will continue asking	4
T-word	... and he will continue asking "why"	1
	total units	24

The process of scoring the complexity of a sentence is the reverse of generating it: the end point of the generation process is the starting point of the analysis and the goal of the analysis is to work backwards to discover which of our rules are manifested in the sentence.

B. The Coding Sheet

The results of the analysis are recorded on a coding sheet which is divided into seven vertical columns. The headings of the columns are as follows:

Sent: The number or letter used to identify the sentence being coded.
K or T: Kernel or Transformed Sentence
M: The number of units (obligatory + optional) in the matrix.
1b: Single-base Transformations
2b: Double-base Transformations
K opt: Optional units in the insert kernel (units may be scored here only when double-base transformations are used)
Tot: Total units for the sentence.

Figure I is a sample coding sheet on which the three examples from the preceding section have been scored. Note that the number of units scored for each transformation is entered alongside the name of the transformation. Note also that in example (b)
is added by the second application of T-conj-P. Since T-conj-P is a single-base transformation, this unit cannot be placed in the K opt column and must be scored as part of the matrix under M.

FIGURE I. SAMPLE CODING SHEET

Sent.	K or T	M	1B	2B	K opt	Tot
(a)	T	2	indir obj-1 neg-1 dc-1 impera-2 wh NP-2	indirq-4		14
(b)	T	2+2	why-2 yes/no-1 do-1 neg-1 conjp-2 conj p-2			13
(c)	T	2+1	transp adv-1 obj-1 neg-1 do-1 word-1	conjK-3 conjK-3 conjK-3 ing inf-4	+2 +1	24

The following information is recorded for each subject on a summary sheet (two summary sheets for each session: one for mothers and one for children):

1. Name (or code number) of mother or child
2. Number of kernel sentences
3. Number of transformed sentences
4. Total number of single-base transformations used
5. Mean number of single-base transformations per sentence
6. Total number of double-base transformations used
7. Mean number of double-base transformations per sentence
8. Mean number of units per sentence
9. Number of different transformations used: single-base, double-base and total.

C. Coding Deviant Sentences

The following types of sentences are not coded:
(a) One-word answers to questions.
(b) Sentences ending with three dots (...).
(c) Sentences containing blanks (_____), unless the part of speech is obvious.
(d) Sentences with (...) in the middle, unless there is no significant interference.

Otherwise, nearly all sentences, even though deviant, can be coded. It should be noted that the term "deviant" is used here to describe utterances which deviate from our set of rules and in no sense is it meant to stigmatize such utterances. In some cases, the deviations indicate shortcomings in the rules; in others, they reflect common colloquial or dialectical usage (e.g. I'm gonna, I wanna, I ain't); but often they also result from the incomplete or incorrect application of the rules. The goal, then, is to reduce the complexity scores of the last type but not that of the first two types.

The usual method for indicating deviant sentences on the coding sheet is to place an asterisk next to the name of the transformation and, when appropriate, to decrease the number of units for that transformation. Thus the sentence

Is this bigger or this one?

would have the entry conj K* 2

on the coding sheet, indicating that one unit had been subtracted for the missing verb. This assumes that the complete sentence would be

Is this bigger or is this one?

But the coding of the sentence I ain't going

would include neg*-1 with no decrease in the number of units indicating the complete but non-standard application of the rule.

It often happens that the deviation is indicated implicitly, as when one transformation is entirely omitted from a normal sequence of transformations as in What those are?
Where T-wh-NP should have been accompanied by T-yes/no, but was not. The reverse case also occurs: an extra transformation may be added to a sequence of rules, as in I know what's that. where T-yes/no has been added to the normal sequence T-wh-NP, T-indirq. In these and similar cases, all the transformations which have been applied are scored and there is no special indication of deviation. The general principle, then, is that deviation is indicated by an asterisk only when the operations performed by a given transformation are carried out incompletely or incorrectly, and not when a transformation is applied correctly when it should not be applied at all, nor when a transformation is omitted altogether.

If the deviation occurs in optional elements in the matrix which only get one unit, the procedure is to award the unit but to place an asterisk next to it, e.g.

(a) The man smoking.
(b) I gonna get some

would be coded

(a) 2 + 1*
(b) 2 + 1s*,

in the matrix to indicate the omission of "be" in each case, i.e. "ing" and "going to" were added, instead of "be + ing" and "be going to".

If the VP or NP, or the "essential part" thereof (i.e. the Verb or Noun/Pronoun) is omitted, this is indicated simply by not scoring any units for the missing element, e.g. That a ball. Ball. each receive only one unit in the matrix for an NP, and none for the missing verb.

Similarly, if the NP and VP are missing completely, and only optional or transformational elements are present, it is obvious that no units can be scored for the NP or VP.

e.g. (a) blue.
(b) because it does
(c) why?

are scored as follows:

(a) Matrix = + 1
(b) Matrix = 0, T-sub 2
(c) Matrix = + 1, T-why, T-?

But if an NP is the only element present, even though it is obviously part of a VP, it is scored 1 unit, e.g. Too heavy (M = 1 + 1).
For sentences of the type: That go there? Where he go?
where the element "tense" is missing, see the discussion under T-yes/no section IIIB.

If the verb is missing but, in addition to an NP, there is an optional element present which can be either optional or obligatory depending on what verb might be present (e.g. an adv-p. gets + 1 after a verbal but not after "be"), the element is scored as optional, e.g. Car on the track
receives 1 + 1 units in the matrix. It will be noted that this is an exception to the principle stated earlier, that the simplest derivation is assumed. Following is a list of the major types of deviant sentences with examples from our data. (Scoring of the relevant features only is indicated.)

1. NP missing:
   (a) Take this out? M = 1 (Cannot be an imperative because of the question mark).
   (b) Guess what? M = 1 (But "guess what" is scored as an imperative)

2. Verb missing.
   (a) What this door? M = 1
   (b) What about the boy? M = 1 + 1 (Assumes NP + prep-p, Something + about the boy).
   (c) This fire engine. M = 1
   (d) That a doggie. M = 1
   (e) Rene and you and me. M = 1 (plus T-conj-p twice).
   (f) That horsey. M = 1 (It is irrelevant whether the complete sentence would be "That (is a) horsey." or "that horsey (+ VP)"; the scoring is the same in either case.)
   (g) There people in the plane? M = 1 + 1
   (The Matrix is "People in the plane", followed by T-there).
   (h) That mine. M = 1
   (g) This, girl. M = 1 (followed by T-series p; cf. "that girl," M = 1 without T-series p.)
   (h) What the lady name? 1 + 1 (assumes "the lady name + NP", where "lady" gets + 1, not T-pos, since the possessive ending is missing).
   (i) I'll that man and you be doggie. M = 1 + 1 (I that man + will)
   (j) Window up here. M = 1 + 1 (NP + adv-p)

3. NP and VP missing.
   (a) Like Mark and Owen. M = + 1 ("like Mark") plus T-conj-P.
   (b) Why not? M = + 1 (adv-purp), T-why, T-neg, T-?
   (c) Cause it's my birthday. M = Ø, T-sub-2.
   (d) Cause. M = Ø, T-sub-2* = 1 unit

4. Missing Matrix element other than NP or VP (no units subtracted):
   (a) Daddy go. ("tense" missing)
   (b) Jack fell down in water ("art" missing)
   (c) What else you want? ("tense" missing)

5. Incomplete optional element in Matrix.
   (a) You want help me? M = 2 + 1s* (1s* for incomplete aux-s: "want" vice "want to")
   (b) Want to go bathroom? M = 1 + 1s + 1* (1 for go, + 1s for "want to", + 1* for "bathroom" with "to the" missing.)
   (c) I'm think about these, M = 2 + 1* (+ 1* for "be"(=am) without "ing")
   (d) The train coming. M = 2 + 1* (+1* for "+ing" without "be").
   (e) I got to sit on it. M = 2 + 1s* (1s* for "got to" vice "have got to" or "have to").
   (f) You try and open the door. M = 2 + 1s* (+1s* for "try and" vice "try to")
   (g) Piggie try eat. M = 2 + 1s* (+ 1s* for "try" vice "try to")
   (h) You gonna do it. M = 2 + 1s* (+ 1s* for "going to" (= gonna) vice "be going to")

6. Transformation applied incorrectly or incompletely:
   (a) Let's see do it balance. T-yes/no, T-do, T-imper-b, T-sub b * - 3 units.
   (b) You want it to. T-VTO-b* - 1 unit
   (c) Mommy, what do that says? T-yes/no* - 1 unit (no units can be subtracted since the transformation only receives one unit).
   (d) Do Debbie be coming soon? T-yes/no* - 1 unit (cf. example c)

7. Transformation not applied. What those are? (T-yes/no wasn't applied).

8. Extra transformation applied:
   I don't know what's this. (T-yes/no should not have been applied).

9. Colloquial or dialectical variants (scored like standard sentences):
   (a) C'mere = come here
   (b) What'cha have? = What do you have?
   (c) I dunno = I don't know.

(d) There go some puzzles up there = There are some puzzles up there.
(e) Where's this go? = Where does this go?
(f) It is getting fixed = It is being fixed.
(g) How come? = why?

D. V.E. SAMPLE SCORING

This section contains a number of sentences from the transcript of K.D., age 4. The left-hand column contains the remarks of an observer and should be disregarded. The right-hand column contains the utterances of the mother and child (indicated by M; and C;, respectively) which are to be coded. When the subjects' utterances are preceded by two numbers, the second number identifies the sentence and is the one that is recorded on the coding sheet (the first number identifies the corresponding remark by the observer). One hundred sentences (50 for the mother and 50 for the child) beginning with sentence number 213, have been scored on the coding sheets following the transcript. Note that there are separate sets of sheets for the mother and the child. Note also that if a sentence number on the transcript is circled the sentence was not coded, either because it was uncodable (see Section V.E.) or because it was not needed for the several sentences.

The column headings are as follows:
Sent: sentence number
K or T: Kernel or transformed sentence
M: number of units in the matrix
1B: single-base transformations
2B: double-base transformations
K opt: optional units in the insert kernel
Tot: Total complexity score for the sentence.

Items in parentheses were eliminated from the coding.
For an abbreviated list of transformations, see Section VII.
Note that the coding begins approximately one-third of the way through the transcript, and that the first 116 sentences have been omitted.

TRANSCRIPT AND ORIGINAL NARRATION OF K.D., AGE 4

213.	Says that he found one.	213-216.	C: I found one.(simult)
214.	The mother labels it T.	214-217.	M: That's a T.
215.	The mother tells him to find a K that's like his name.	218.	C: A T
		219.	M: Find a K
		215-220.	M: See if you can find a K that looks like your name
216.	He finds one.	216-221.	C: Here.
217.	He is correct	218-222.	M: That's a K?
218.	But the mother isn't sure about it.	223.	C: Yeah.
		224.	M: You sure?
219.	Then she labels it K.	219-225.	M: That K?
		226.	M: Where is it?
		227.	C: Noooo.
220.	The mother asks him if he wants to spell his name.	220-228.	M: You want to spell your name?
		229.	M: Spell your last name then.
		230.	C: What's this up here?
221.	Kevin says he wants to play outside.	221-231.	C: Ma, I'm going...
		232.	C: I wantta play out there.
222.	The mother asks him out where.	222-233.	M: Out where?
223.	He points outside.	223-234.	C: Out there.
		235.	M: That's the hallway.
224.	Kevin walks over to the magnet board.	225-236.	C: Mommy what's this?
225.	Requests a label.		
226.	The mother labels it board.	226-237.	M: A board.
227.	Kevin says he wants to draw.	238.	C: A board?
		239.	C: To draw on?
228.	The mother says she doesn't have any apaper.	240.	M: I don't know.
		241.	M: I think you can.
229.	Kevin points to the magnet board.	227-242.	C: I want to draw .
230.	Says that he'll draw on that.	228-243.	M: I don't have any paper.
		230-244.	C: I draw on here.
231.	The mother informs him they didn't have any chalk.	231-245.	M: I don't have any chalk.
232.	Kevin says that she should have brought some chald with her.	232-246.	C: You should have brought some chalk with you.
		247.	M: We can draw when we go home.
		248.	M: OK.
		249.	C: (noise)
233.	Kevin asks "on my board?"	233-250.	C: Draw on my board.
		251.	M: No.
234.	The mother informs him, on his paper.	234-252.	M: On the paper.
235.	He asks which paper?	235-253.	C: What paper?

236.	She says in her book.	236-254. M: In your book.
237.	Kevin says he doesn't have any.	237-255. C: I don't have any.
238.	The mother corrects him.	238-256. M: You do so have a book ___.
239.	Negates it.	241-257. C: No, scribble, scrabble.
240.	Corrects him.	242-258. M: Scribble scrabble?
241.	Saying that he does.	259. C: Yeah.
242.	They talk about his writing in the book and about scribbling.	260. M: You scribble scrabble in it too.
		261. C: OK
243.	Kevin picks up an I	262. C: (noises)
244.	His mother requests a label.	263. C: Found one.
		244-264. M: What's that?
245.	He labels it "one".	245-265. C: A one.
246.	His mother negates this.	
247.	Corrects him, saying that it's an I.	247-266. M: That's a I.
248.	Kevin insists that it's a one.	248-267. C: uh uh? (i.e. no)
		268. M: That's a I.
		269. C: I.
		270. M: Here.
		271. M: Spell your name.
249.	And his mother keeps labeling it "I".	250-272. M: K-E.
250.	Kevins' mother puts the K and the E down.	273. M: Got to find a V.
		274. C: Ah, here's a one.
		275. M: That's a I.
251.	Kevin's mother directs him to spell his name.	276. M: This is a I.
		277. C: I know you say it's not but it is.
252.	Kevin is looking at the mirror.	253-278. C: I want to see somebody.
253.	Saying that he wants to see someone.	279. M: You want to see somebody?
		280. C: Yeah.
254.	The mother asks why.	254-281. M: Why?
255.	Kevin turns to the glass.	255-282. C: From this glass.
256.	Says that he wants to see somebody.	256-283. C: I want to see somebody.
257.	Kevin continues to say that he wants to see somebody.	284. M: Whatcha want to tell them when you see them?
		257-285. C: I can just tell them "I want to see somebody."
		286. M: (laughs)
		287. C: I'm gonna try.
		288. M: Try?
		289. C: Try to see somebody.
258.	The mother says that he can see himself.	259-290. M: You can see yourself, look.
259.	If he looks.	259-291. C: No.
260.	Kevin says he wants to see someone else in another school.	260-292. C: I wanna see somebody else in the school.

261. Kevin's mother says that he used to tell her about when he went to school here.

262. And he used to play a game with a lady in _____

263. Kevin says that he saw someone else in the glass.

264. They had the light on.

261-293. M: You remember when you used to come here and they used to put you in the room?

294. C: Yeah.

262-295. M: You and the teacher?

296. C: Yes

297. M: Well you didn't see any body then, did you?

298. C: Unhuh.

263-299. C: I saw a lady.

300. M: You did?

301. C: From that glass.

302. M: You did?

303. C: unhuh.

264-304. C: Cause all the light was off.

305. M: When we put the lights out, it's gonna be dark in here.

K.D., Age 4 Syntactic Complexity Coding Sheet

Sent.	K or T	M	1B	2B	K Opt	Tot
230	T	2+1	wh-NP-2			6
			?-1			
234	K	+1				1
236	T	2	voc-1			
			wh-NP-2			
			?-1			6
238	T	1	?-1			2
239	T		?-1	to adja*		
				-4		5
242	K	2+1s				3
244	K	2+1				3
246	K	2+3				5
250	K	1+1				2
253	T	1	wh art-3			
			?-1			5
255	T	2	neg-1			
			do-1			4
257	T	1	interj-1			2
261	T		interj-1			1
263	K	1				1
265	K	1				1

Mother of K.D. Syntactic Complexity Coding Sheet

Sent.	K or T	M	1B	2B	K Opt	Tot
213	K	1				1
214	T	2	impera-2			4
215	T	2	impera-2	sub b-5	+1	10
217	K	2				2
219	T	2	impera-2			4
220	T	2	impera-2	rel a-4		
				sub b-5		14
222	T	2	?-1			3
224	T	1	?-1			2
225	T	1	?-1			2
226	T	2	wh advp-2			
			yes/no-1			
			?-1			6
228	T	2+1s	?-1			4
229	T	2+2	impera-2			6
233	T	+1	wh advp-2			
			?-1			4
235	K	2				2
237	K	1				1
240	T	2	Neg-1			
			do-1			4
241	T	2		suba*-3	+1	6
243	T	2+1	Neg-1			
			do-1			5
245	T	2+1	Neg-1			
			do-1			5
247	T	2+1		sub 2-4	+1	8
248	T		interj-1			1
251	T		interj-1			1
252	K	+1				1
254	K	+1				1
258	T	1	?-1			2
260	K	2+2				4
264	T	2	wh NP-2			
			?-1			5
266	K	2				2

VI. THE CONCEPT OF CRITICAL STRUCTURE

For scoring of transformation imitation tasks an approach involving estimation of the operational essence, or "critical structure" of a transformation has been attempted. When a transformation is written in terms of the highest* level components permitting a definition of the transformational rule, the "critical structure" consists of a part of each element contiguous to the transformational operation(s) plus any new element required by the transformation. For example, the transformation T-negative can be written NP + tense - modal + verbal + X --> NP + tense - modal + not + verbal as: John is going to the store --> John is not going to the store. By the rule, the "critical structure" would consist of is + not + go(ing).

With double-base transformations it is further required that at least part ** of each NP and VP present in the ("result") final sentence be counted in the critical structure. Thus in the transformation T-conj-K (NP_1 + VP_1 + NP_2 + VP_2 --> NP_1 + VP_1 + conj + NP_2 + VP_2), the critical structure must have elements of NP_1 and VP_2 as well as those satisfying the above conditions. E.G. John walked and they followed would be a satisfactory 'critical structure" of the sentence "John from next door walked to the beach and they followed him closely in a black sedan"**

---

*
When written as in this manual, thus S being higher than NP_1 which is higher than N

**
So would, e.g., Next door to the beach and they closely, Since part as used above has not been restricted. Since on the whole, subjects don't give such poetic responses on imitation tasks, an arbitrary group of acceptable definitions of part has not been worked out.

VII. LIST OF TRANSFORMATIONS

Kernels:

1.	Noun phrase kernel	1 unit
2.	Noun phrase optionals	1 unit each
3.	Verb phrase kernel	1 unit
4.	Verb phrase optionals	1 unit each

Single-Base Transformations:

		Frank #	Units	Examples
1.	T-adv-sm	(a-16)	1	Fortunately, he left.
2.	T-affirm	(5)	1	I HAVE spoken.
3.	T-conj-p	(17)	2	We eat fish and cheese.
4.	T-compar-a	(A-5a)	1	I am bigger.
5.	T-do	(2)	1	I do think you should go.
6.	T-go (come)	(A-17)	1	He will go work
7.	T-imper-a	(18-a)	2	Look!
8.	T-imper-b	(18b)	3	Let's look.
9.	T-imper-c	-	4	Let me see.
10.	T-imperf	(a-2)	2	He used to do that.
11.	T-indir. obj.	(10)	1	We told you the story.
12.	T-interj.	(12)	1	Wow, it's late.
13.	T-invers	(8)	1	Well, it was this way.
14.	T-neg	(4)	1	We are not ready.
15.	T-obj.	(6)	1	I saw him.
16.	T-passive-b	(20a)	4	The fly was killed.
17.	T-passive-b	(20b)	5	The fly was killed by me.
18.	T-passive-inf.a	(A-4a)	4	The fly has to be killed.
19.	T-passive-inf.b	(A-4b)	5	The fly has to be killed by me.
20.	T-? (q)	(3)	1	You know him?
21.	T-refl.	(A-1)	2	He hit himself.
22.	T-same-as-a	(A-6a)	2	It was the same as this one.
23.	T-series-p	(11)	1	John, Sam went.
24.	T-superl.a	(A-7a)	1	I am most beautiful.
25.	T-superl.b	(A-7b)	2	I am the most beautiful.
26.	T-there	(16)	2	There will be a dance.
27.	T-transp.-adv	(A-11)	1	I frequently read books.
28.	T-transp.-clause	(A-12)	1	Why he did it, I don't know.
29.	T-transp.-NP	(A-13)	1	My bike I like to ride.
30.	T-transp-preart	(A-10)	1	The children are all going.

31.	T-voc	(9)	1	John, the tea is ready.
32.	T-vt	(7)	1	I put them away.
33.	T-wh-adv-m	(14b)	2	How
34.	T-wh-adv-no	(?)	2	How many
35.	T-wh-adv-p	(13)	2	Where
36.	T-wh-adv-t	(14a)	2	When
37.	T-wh-art	(19)	3	Which (man)
38.	T-wh-intensifier	(a-8)	3	How pretty that is.
39.	T-wh-NP	(15)	2	Who, what
40.	T-wh-possessive	(A-5)	4	Whose hat this is.
41.	T-why	(14a)	2	Why
42.	T-word	(?)	1	"John" is a noun.
43.	T-yes/no	(1)	1	Was he going there
44.	T-yes/no 2A	(1b)	3	They are devils, aren't they
45.	T-yes/no 2b	(1b)	2	They don't speak well, do they

Double-Base Transformations:

46.	T-as	(A-6a)	5	I am as old as he is.
47.	T-apos	(23)	3	I see my sister, Jane, today.
48.	T-conj-K	(22)	3	John goes and Tom stays.
49.	T-compar.b	(?)	5	I am bigger than he is.
50.	T-indir.q	(26)	4	I saw what they did.
51.	T-indir.q.b	(26a)	5	I saw what to build.
52.	T-ing-a	(36a)	5	We noticed the shouting.
53.	T-ing-b	(36b)	7	We noticed Bob's shouting
54.	T-ing-c	(36c)	6	We noticed Bob shouting.
55.	T-ing-adj.	(36d)	4	I go upstairs laughing.
56.	T-ing-inf.	(B-5)	4	He started laughing.
57.	T-parenth	(25)	4	Yesterday (I worked) I fainted
58.	T-pos	(32)	5	I want the general's books.
59.	T-rel-a	(31)	4	We saw the birds which sang.
60.	T-rel-b	(31a)	3	Show me the one you made.
61.	T-rel-when	(31b)	4	I knew the day when you came.
62.	T-rel-where	(31c)	4	I know the spot where you went.
63.	T-same-as-b	(?)	4	It was the same as I thought.
64.	T-same-as-c	(?)	5	" as when I had thought.
65.	T-sep	(34)	5	They said "Come in."
66.	T-series-k	(21)	2	John goes, Sam stays.
67.	T-sub-a	(29a)	4	He knew they would win.

68.	T-sub-b	(29b)	5	He knew that they would win.
69.	T-sub-c	(?)	4	The story is that X.
70.	T-sub-d	(B-6)	4	It was so good that X.
71.	T-sub-2	(30)	4	Whenever I go, I see him.
72.	T-subj.a	(?)		
73.	T-subj.b			
74.	T-subj.c			
75.	T-to-a	(35a)	5	It will be necessary to go
76.	T-to-b	(35b)	7	It will be necessary for him to go.
77.	T-to-adj-a	(B-2a)	5	He makes hats to look at.
78.	T-to-adj-b	(B-2b)	5	He makes hats for them to see.
79.	T-to-adv	(?)	4	He is ready to go.
80.	T-to-NP-a-subj	(B-1a)	5	To die for the state is brave.
81.	T-to-NP-b objt	(B-1b)	4	He wants to go to school.
82.	T-to-NP-c pred.nom	(B-1c)	4	The aim is to hit the ball.
83.	T-to-purpose-a	(B-3)	5	He does it to earn money.
84.	T-to-purpose-b	(B-3a)	6	He does it in order to X.
85.	T-VTO-a	(28b)	4	They let dance the girls.
86.	T-VTO-b	(29a)	5	They allowed to dance the X.
87.	T-VT3	(27)	4	He thinks pretty the girl
88.	T-want-a	(b-4)	6	I want the man searched.
89.	T-want-b	(b-4a)	8	I want him to be searched.

NOTES:

- These transformations allow spatial latitude: voc, interj., apos, parenth, sep, sub. 2. These allow for two major possibilities: rel, ing-adj. All others require transposition transformations for unnatural positions.

- Things with question marks cannot be imperatives.

Sentences Not Coded For Grammatical Complexity:

1. One word responses to questions
2. Sentences with ...
3. If there was a blank in a sentence and it was obvious what part of speech it was, it was included.
4. Strings of letters, numbers
5. Choo-choo i.e., sounds
6. Cries, whines, laughs, yawns, sneezes.

ABBREVIATIONS AND SYMBOLS

Ad:	Add, addition
Adj:	adjective
Adj-prep.p:	adjectival prepositional phrase
Adv:	adverb, adverbial
adv-freq:	adverb of frequency
adv-m:	adverb of manner
adv-p:	adverb of place
adv-prep.p:	adverbial prepositional phrase
adv-purp:	adverb of purpose
adv-sm:	adverbial sentence modifier
adv-t:	adverb of time
art:	article
aux:	auxiliary
aux-s:	S-auxiliary
comp:	complement (a dummy symbol)
conj:	conjunction
def:	definite article
demon:	demonstrative
det:	determiner
indef:	indefinite pronoun
infl:	inflection
insert:	insert kernel
int:	intensifier
K:	kernel
M:	matrix
M:	modal
N:	noun
Nondef:	nondefinite article
NP:	Noun Phrase
obj:	object
part:	participle
Pers:	personal pronoun
pre-art:	pre-article
prep:	preposition
pron:	pronoun
pron-indef:	indefinite pronoun
Pron - pers:	personal pronoun
pron - poss:	possessive pronoun
prt:	particle
P-rule :	Phrase Structure Rule
s:	sentence
Sent:	Sentence
sub:	subtract, subtraction
sub:	subordinating conjunction
substantive:	a noun or an adjective
T:	transformation, transformational rule
T:	transformed sentence
tr, trans:	transpose, transposition
T-rule:	Transformational Rule

V:	verb, including <u>verbal</u> and <u>be</u>
Vb:	"become" - type verbal
Verbal:	a verb other than <u>be</u> or <u>aux</u>
Vh:	"have" - type verbal
VI:	intransitive verbal
VP:	Verb phrase
Vs:	"seem" - type verbal
Vt:	transitive verbal
1BT:	single-base transformation
2BT:	double-base transformation
-->	may be rewritten as
≠	is not equal to
∅	zero
[ ]	an optional element
{ }	obligatory selection of an element within the brackets

BIBLIOGRAPHY

I. REFERENCES EMPLOYED IN FORMING THE ANALYSIS SYSTEM

Chomsky, N. (1957) Syntactic Structures (The Hague)

Chomsky, N. (1965) Aspects of the Theory of Syntax (Cambridge)

Frank, S. and Osser, H. (1970) A Psycholinguistic Model of Syntactic Complexity, Language & Speech, (in press)

Menyuk, P. (1964) Syntactic Rules used by Children from Preschool through first grade, Child Development, 35, 533

Roberts, P. (1964) English Syntax (New York)

II. REFERENCES EMPLOYING ELEMENTS OF THE PRESENT SYSTEM

Baldwin, A.L. and Frank, S.M. (1969) Syntactic Complexity in Mother-Child Interactions (Mss, presented to Society for Research in Child Development Santa Monica, March, 1969)

Blank, M. and Frank, S.M. (1969), Story Recall as a Measure of Psycholinguistic Performance (Mss.)

Osser, H.,and Wang, M., and Zaid, F., (1970) The Young Child's Ability to Imitate and Comprehend Speech: Comparison of Two sub cultural groups Child Development,(In press)

Appendix B *

LANAL-I:

An Automated Content-Analysis System

Introduction

LANAL-I is an extremely flexible program for the analysis of written records. It was developed as a tool for the content analysis of narrative reports of naturalistic observations. LANAL-I accepts as input a "transcript", that is, a series of sentences. By means of dictionary and a rule system, it maps this series of sentences into an output. The dictionary and the rule system, and thus the form of the output, are under the control of the user, and in fact, form part of the input to the program. Some possible types of operations which may be performed are; grammatical transformations of input sentences; syntactic parsing of input sentences; or conceptual analysis of a series of sentences.

Words as "Feature-Clusters"

Fundamental to the operation of the program is the idea of handling words as clusters of syntactic and/or semantic features. This treatment has much in common both with the treatment of syntactic and semantic features by Chomsky (1965), and with the replacement of word stems by concept numbers in the information retrieval system of Salton (1968), although it is not identical with either one. The idea may be best introduced in terms of a simple example. Let us restrict ourselves to the subclass of simple declarative English sentences containing only "the", nouns, verbs, and adjectives. Some examples of such sentences are: "The boy hit the red ball"; "Education produces wise men." Let us further subdivide and restrict our syntactic categories as follows: nouns can be count or noncount, and count nouns can be animate or inanimate; verbs can be present or past, transitive or intransitive; adjectives can be size, color, or personality attribute. Using this (artificial and arbitrary) classification system, we can produce a dictionary by a simple procedure. First, write a word into the left-hand side of a checklist, which is ruled off into columns corresponding to the "features" of the classification system outlined above. Second, place a check in the columns which apply to the particular word, and leave the others blank. A "dictionary" of this type is shown in Table 1.

---

*

This appendix was prepared by Paul Ward.

Table 1

A Simple "Checklist" Dictionary

WORD	NOUN	VERB	ADJECT	COUNT	NONCOUNT	ANIMATE	INANIM.	PRES.	PAST	TRANS.	INTRANS.	SIZE	COLOR	PERS. ATTR.
BALL	X			X			X							
BOY	X			X		X								
BLUE			X										X	
EDUCATION	X				X									
HITS		X						X		X				
LOOKED		X							X		X			
MEN	X			X		X								
PRODUCES		X						X		X				
RED			X										X	
WISE			X											X

Given a dictionary of this form, it is a simple matter to indicate the definition of a word. The definition is simply some representation of the checklist, showing which features have been checked off and which have not. A particularly simple representation is in terms of a binary number, using "zero" for "not checked" and "one" for "checked" and letting the successive digits represent successive items on the checklist. Using this formalism and referring to Table 1, the definition of "education" is 10001000000000, and the definition of "looked" is 01000000101000. This binary number notation does not imply any assumption about classifying words in terms of "bits" of information; it is simply a convenient form for classification and processing.

There are a number of observations to be made about Table 1 which have important implications for the rest of the analysis scheme. One of the most obvious observations is that there is a large amount of redundancy in the feature system. For one thing, the "noun", "verb", and "adjective" features are actually superfluous, since the features assigned to nouns, verbs, and adjectives form mutually exclusive groups. (This is certainly not true of all classification systems of this type, however. For example, a feature which we may label "inquiry" is common to the verb "asks" and the noun "questioner".) More importantly, features which are at different levels in a hierarchy are treated equivalently. "Count" and "noncount" are subclasses of "noun", and "animate" and "inanimate" are subclasses of "count", but there is nothing in the notation to indicate this. The point is that the "meaning" of a feature, and its relationship to other features, is left entirely to the user's interpretation. In the present case, the marking of nouns might correspond to a "bracketed" notation [e.g. boy (noun(count(animate)))]in which each feature indicates a level of the bracketing. There is also some ambiguity in the meaning of a "zero" entry for a particular word. For the word "red", the zeroes in the "size" and "personality attribute" columns, mean that red is <u>specified negatively</u> for these features. In other words, these features are relevant for the classification of "red" but are found to be <u>absent</u> from "red". On the other hand, the zeroes in "present" and "past" mean that "red" is <u>not specified</u> for these features, i.e., that these features are <u>not relevant</u> for the classification of "red". This distinction is discussed in terms of transformational grammar by Chomsky (1965). The ambiguity in the "zero" specifications could be reduced or eliminated by multiply defining column entries. For example, a "one" in column five could mean "present" if the word was previously marked "verb", and "count" if it was previously marked "noun". However, the rule system to be described below can take account of this problem adequately, and the simple, redundant, linear system described above is probably the most efficient in terms of simplicity of rule-writing.

A final observation to be made is that within a given feature system, two words are "synonyms", or have the same "meaning", just in case the binary numbers specifying their definitions are equal. In

terms of our previous classification system, "blue", "red", "green", and "yellow" are all synonymous, and distinguishing them by their definitions would involve adding extra features to the system.

The Source Dictionary

In order to compile a source dictionary for LANAL-I, a user selects a set of features, (up to 80 are permitted in the present version of the program), and assigns a four letter alphabetic "tag" to each feature. The entries to the dictionary are then coded by the "checklist" method described above. Each word which is expected to be encountered in the course of an analysis should be entered in the dictionary. The actual input to the program consists of a pair of cards for each dictionary entry, the first bearing the word and the second bearing the binary representation of the definition. Following the entries is a list of the "tags," in the same order as they appeared on the checklist. The "tags" are used by the rule system to refer to specific features of a word. Table 2 lists the "tags" assigned to the features from Table 1.

Table 2

Tags Assigned to Features

NOUN - noun

VERB - verb

ADJT - adjective

CNTT - count

NCNT - noncount

ANIM - animate

NANI - non animate

PRES - present tense

PAST - past tense

TRAN - transitive

INTRA - intransitive

SIZE - size

COLR - color

PRAT - personality attribute

"Intrinsic" and "Contextual" Features

In order to facilitate the following discussion, the following notation for representing sentences will be adopted: a sentence is defined as a string of words, or, equivalently, as a string of feature-clusters, and will be written top-to-bottom, with a word at the left side of each line and the associated positively-assigned features of the word (represented by their tags) at the right of the line. Thus the sentence "Education produces wise men" is written as:

EDUCATION	NOUN-NCNT
PRODUCES	VERB-PRES-TRAN
WISE	ADJT-PRAT
MEN	NOUN-CNTT-ANIM

The feature system summarized in Table 2 will now be extended, as follows: the definition of a word shall specify the presence or absence of the features "subject", "direct object", "noun-modifying adjective", "predicate adjective", and "modified by adjective", represented by the tags SUBJ, DOBJ, NMAJ, PRAJ, and MDAJ, respectively. These five features are very different from the initial fourteen. "Education", for example, is NOUN and NCNT in any context in which it occurs, and so may be said to have NOUN and NCNT as parts of its "intrinsic" definition. However, it is SUBJ or DOBJ only in certain contexts, and these features can only be part of its "contextual" definition, i.e., its definition as it is used in a given sentence. Certain features may be intrinsic for some words and contextual for others. "Questioner" is intrinsicly NOUN, but "question" is NOUN in some contexts and VERB in others. "Education produces wise men" may now be rewritten to incorporate the "contextual" features"

EDUCATION	NOUN-NCNT
PRODUCES	VERB-PRES-TRAN
WISE	ADJT-PRAT
MEN	NOUN-CNTT-ANIM
↓	
EDUCATION	NOUN-NCNT-SUBJ
PRODUCES	VERB-PRES-TRAN
WISE	ADJT-PRAT-NMAJ
MEN	NOUN-CNTT-ANIM-DOBJ-MDAJ

The arrow between the two representations indicates a "transformation" or "rewriting" of the sentence, in the sense that the feature-clusters representing the words have been changed. Let us assume that, for a given type of analysis, we are interested only in whether nouns are

modified by adjectives, and not in the particular adjectives used. We may then discard "wise", after using it to mark "men" with MDAJ, and rewrite the sentence as follows:

EDUCATION	NOUN-NCNT-SUBJ
PRODUCES	VERB-PRES-TRAN
WISE	ADJT-PRAT-NMAJ
MEN	NOUN-CNTT-ANTM-DOBJ-MDAJ

EDUCATION	NOUN-NCNT-SUBJ
PRODUCES	VERB-PRES-TRAN
MEN	NOUN-CNTT-ANIM-DOBJ-MDAJ

The adjective "wise" has, in this case, been absorbed by "men", or one might say, has become a (contextual) feature of men. This idea is developed extensively in the generative grammar of Rosenbaum (1967).

The two transformations shown above illustrate the two principal functions of the rule system, namely, to add features to (or delete features from) feature-clusters, and to manipulate strings of feature-clusters by addition, deletion, or reordering. The formalism for the rule system is the subject of the next section.

It is worthwhile to emphasize once again that a "feature" of a word has no fixed meaning except as interpreted by the user of the program. In particular, in the course of syntactic analysis, contextual features may be added to a word to show that it dominates a node (e.g., DNPP- dominates Noun Phrase) or is dominated by a node (e.g., NPPD-dominated by noun phrase). The ultimate test of the usefulness of a program such as this for a particular application, is whether a suitable notation can be invented within the basic constraints of the system.

The "Rule Shorthand" for LANAL-I

The basic unit of the "rule shorthand" is the transformation or rewriting rule. The most general representation for this rule is simply A=B, meaning "rewrite context A as context B," or "if a portion of the transcript is encountered which meets the specifications A, rewrite it as specified by B". Ending the rule with a period means that it is to be applied just once to each sentence. If it is ended with a question mark, it becomes a "recursive" rule, which is applied repeatedly to a single sentence until it becomes inapplicable. Recursive rules are useful for locating constructions which may appear more than once in a sentence. The "A" part of the rule can contain two types

of subunits: A1 (:A2) = B., where the parentheses indicate that the A2 subunit is optional. The meaning of A1:A2 is "If context A1 is encountered in the sentence currently being processed, and context A2 is encountered in the preceeding sentence." There is also an alternate form A1 (::A2) meaning "If context A1 is encountered in the current sentence, and context A2 is encountered in one of the five preceeding sentences."

Each of the three subunits A1, A2, and B consist of strings of elements separated by "+":

A11 + A12 + ... + A1n (:A21 + A22 + ... A2m = B1 + B2+ ...+ Bp.

There are two basic types of elements; word designators and string designators. A string designator is simply one of the symbols (X1, X2, X3, X4, X5, X6). A word designator consists of a number of sub-elements separated by hyphens. These subelements may be: word numbers, which are simply numerals enclosed in parentheses, such as (1) or (4); feature tags, such as NOUN or SUBJ, which may be preceeded by the negation symbol ¬ ; literals, which are actual words enclosed in quotation marks, such as "PRODUCES" or "WISE"; and "internal comparisons" which are members of the set ($1, $2, $3). String designators which appear to the right of the "equals" sign may be followed by feature tags (e.g., X1 - NPHR). This provides a way of assigning a feature to a whole string of words. Some examples of word designator elements are (3), (1) - NOUN - ANIM (2) - "THE" - $2.

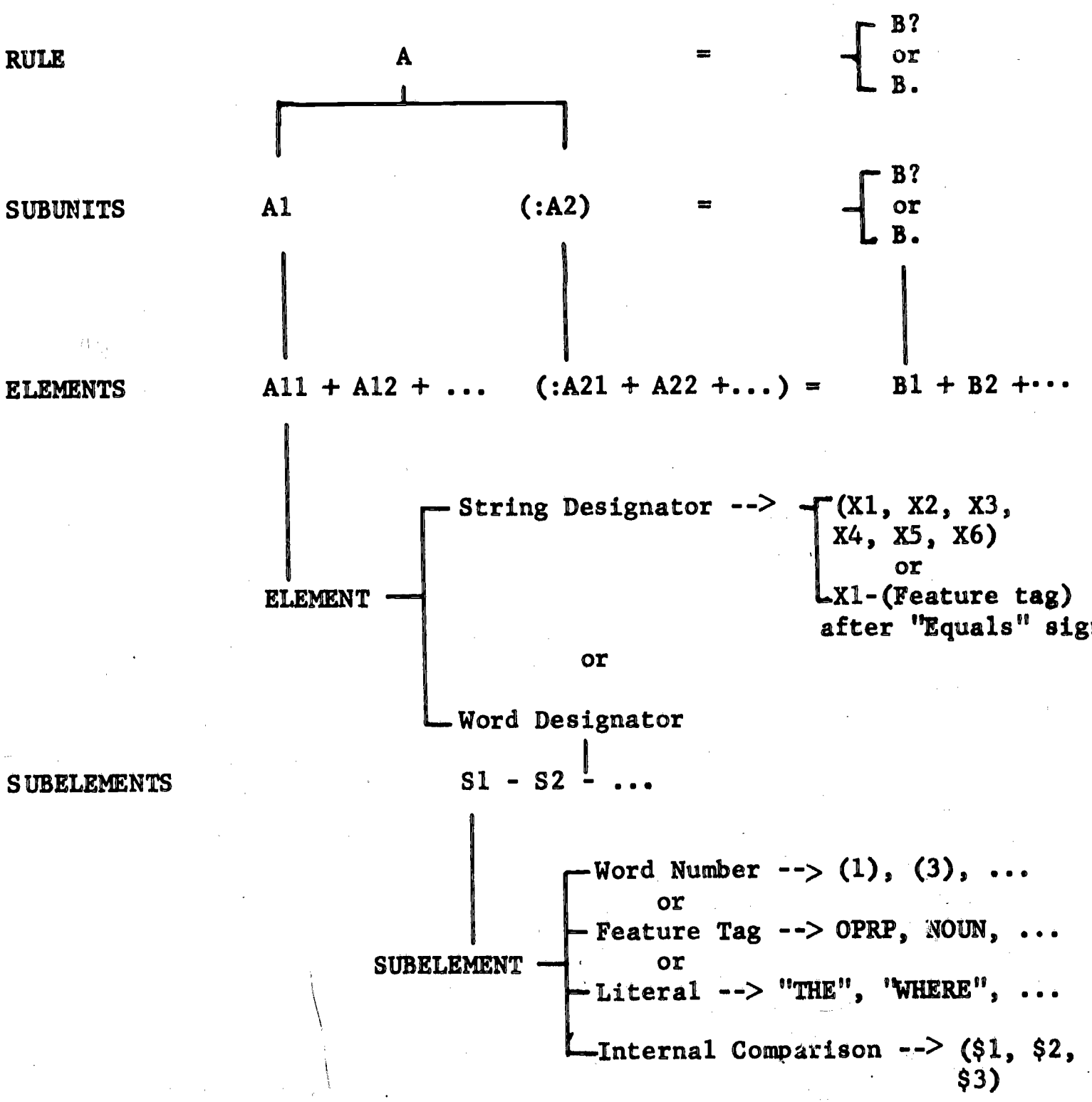
The following example illustrates the writing of rules for a specific part of an analysis. Suppose we are analyzing sentences which contain prepositional phrases such as "to college", "in the cellar", or "on the green sofa". We wish to mark the nouns which are objects of prepositions with the feature OPRP. Our system contains the features VERB, ADVB (adverb), ARTC (article), NOUN, ADJT, and PREP (preposition) as intrinsic features of the word definitions. To handle nouns occuring in phrases such as "to college", which contain only a preposition and a noun, we write the following rule:

X1 + (1) - PREP + (2)- NOUN + X2= X1 + (1) +(2) - OPRP+X2.

The English paraphrase of this rule is "If the sentence currently being processed consists of a string of words X1 (including the null string), followed by a word marked PREP, followed by a word marked NOUN, followed by another string of words X2, rewrite it as follows: copy the string X1, then the word marked PREP, then the word marked NOUN with the added feature OPRP, then the string X 2."

Figure 1

Structure of the LANAL-1 Rule System



this rule, applied to the sentence "The plane flew to Boston yesterday", transforms it as follows:

THE	ARTC
PLANE	NOUN
FLEW	VERB
TO	PREP
BOSTON	NOUN
YESTERDAY	ADVB

↓

THE	ARTC
PLANE	NOUN
FLEW	VERB
TO	PREP
BOSTON	NOUN-OPRP
YESTERDAY	ADVB

In this particular application of the rule X1 designates "the plane flew", (1)-PREP designates "to", (2)-NOUN designates "Boston", and X2 designates "yesterday". The designator X1 in the combination X1+(1)-PREP Actually indicates "any string whose members do not contain the feature PREP." Similarly, X2 in the combination (2)-NOUN+X2 = indicates, "Anything else in the sentence following the occurrence of a word marked NOUN". The rule (1)-PREP + (2)-NOUN + X1 = (1) + (2) - OPRP + X1, handles only <u>initial</u> prepositional phrases, while the rule X1 + (1) - PREP + (2) - NOUN = X1 + (1) + (2) - OPRP, handles only <u>terminal</u> phrases. The precise meaning of the designator (2)-NOUN is "any word with the feature NOUN marked positively, <u>irrespective of any other features</u>."Similarly, the designator (1)-NOUN - OPRP means "any word specified positively for NOUN and negatively for OPRP, irrespective of its specification for any other feature." This property of the rule system avoids the problem mentioned earlier of the ambiguity between "negatively specified" and "not specified", since only the features indicated by the word designator are examined by the program in its search for contexts, and irrelevant features can be ignored.

A version of the rule which would mark the noun with OPRP and <u>delete</u> the preposition is:

X1 + (1) - PREP + (2) - NOUN + X2 = X1 + (2) - OPRP + X2.

This rule transforms"Bill went to Rome" as follows:

BILL	NOUN
WENT	VERB
TO	PREP
ROME	NOUN

↓

BILL	NOUN
WENT	VERB
ROME	NOUN-OPRP

In this example the string X2 is null.

A slightly more complex rule can be written to handle phrases such as "up the high hill." Two possible rules of this type are:

(a) X1 + (1) - PREP + (2) - "THE" + (3) - ADJT + (4) - NOUN + X2 =
X1 + (1) + (2) + (3) + (4) - OPRP + X2

(b) X1 + (1) - PREP + (2) - ARTC + (3) - ADJT + (4) - NOUN + X2 =
X1 + (1) + (2) + (3) + (4) - OPRP + X2

Rule (a) illustrates the use of the literal word reference within a word designator element. The difference between rules (a) and (b) is one of specificity. Rule (a) will only handle prepositional phrases containing"the". Rule (b) will handle phrases containing any word specified positively for ARTC, and is adequate for phrases such as "through an open door" or "by a beautiful woman", provided "a" and "an" are in the dictionary and are marked ARTC.

All of the rule examples given so far have used only A1 elements, since they have made no references to inter-sentence connections, i.e., to "hyper-syntax" as defined by Wooley (1966). The following example makes use of an A2 element, and also of an "internal comparison" within a word designator element. Consider the two sentences, "The child asks where the toy is," and "The mother ignores the child's question." In the second sentence the word "question" stands for or refers to the first sentence, just as an ordinary pronoun stands for or refers to the occurrence of a noun. In more general terms, the occurrence of "child's question" in a coherent series of sentences presupposes the previous occurrence of some sentence like "the child asks ------." Let us assume that in second sentence "child's" has been transformed to "child" and marked POSS (possessive-this is done in the initial phase of processing the transcript), and that both "asks" and "question" are marked with the feature INQR (inquiry). The following rule will replace "question" by the sentence to which it refers:

X1 + (1) - NOUN - POSS - $1 + (2) - NOUN - INQR: (3) - ARTC + (4) - $1 + (5) - VERB - INQR + X2 = X1 + (1) - ⌐ POSS + "WHEN" + (3) + (4) + (5) + X2

This rule will transform the sentences mentioned above, as follows:

THE	ARTC
CHILD	NOUN
ASKS	VERB-INQR
WHERE	RLPR
THE	ARTC
TOY	NOUN
IS	VERB
THE	ARTC
MOTHER	NOUN
IGNORES	VERB
THE	ARTC
CHILD	NOUN-POSS
QUESTION	NOUN-INQR

↓

THE	ARTC
MOTHER	NOUN
IGNORES	VERB
THE	ARTC
CHILD	NOUN
WHEN	CONJ
THE	ARTC
CHILD	NOUN
ASKS	VERB-INQR
WHERE	RLPR
THE	ARTC
TOY	NOUN
IS	VERB

The "child asks -----" sentence is not changed by the application of this rule. A2 elements are used in context identification, and parts of them may be copied into the B section of the rule, but only the sentence mentioned in the A1 element is changed. Note the application of the internal comparison operators; they always occur in pairs, and indicate the occurrence of the same word in two different places without indicating a specific word. This adds a great deal of generality to a rule; the rule written above will handle, "the mother asks ----", and "---- the mother's question", or any similar pair of sentences. It will also handle any noun-verb pair having the feature INQR, and thus can treat synonyms such as "Query" for "question" without trouble. Note also that the (1) - ⌐ POSS designator in the B section of the rule has deleted the feature POSS from "child" in the transformed sentence.

The program applies the set of rules to the sentences sequentially with each rule operating on the sentence as transformed by the last rule. This adds a great deal of power to the rule system, since features introduced in an earlier part of the analysis may be used as cues for later rules. It also means that references to previous sentences apply to those sentences as transformed by the rules.

In the discussion of the checklist dictionary, the problem of ignoring hierarchical order in the feature system was raised. Now that the rule system has been introduced, it should be pointed out that rules can be written to apply to any given level of a hierarchy. Consider the series of word-designators (1)-"BOY", (1)-ANIM, (1)-COUNT, (1)-NOUN. The first designator recognizes only the specific word "boy", the second any animate noun, the third any count noun, and the last any noun at all. By writing the rule system appropriately, hierarchical distinctions can be maintained, and, more importantly, flexibility can be introduced into a hierarchy if necessary.

Although it was developed independently, the rule-writing shorthand described here bears a number of resemblances to the Reduction Analysis rule-writing system in IBM's Recognition Grammar, developed by Culicover et.al. (1969). The RA rule system, however, at least in its present form, is tied strongly to a linguistically oriented, tree-structure analysis of single sentences. The system described here, while less sophisticated in linguistic terms, provides somewhat more flexibility in terms of the form of the analysis, and in particular allows for inter-sentence operations.

Further Examples of Rewriting Rules

(a) Roberts (1964) gives an example of a double-base transformation which he calls T-POS. This will rewrite the pair of sentences "John has a car" and "I bought the car" as "I bought John's car". Assuming that our feature system has an entry ATTR (attribution) which is common to verbs such as "has", "owns", "possesses", etc.

The rule to apply T-POS is written:

(1) - NOUN + (2) - VERB + (3) - "THE" + (4) - NOUN - $1: (5) - NOUN + (6) - VERB - ATTR + (7) - "A" + (8) - $1 = (1) + (2) + (5) - POSS + (8).

This rule will rewrite the above sentences as follows:

JOHN	NOUN
HAS	VERB-ATTR
A	--
CAR	NOUN

I	NOUN
BOUGHT	VERB
THE	--
CAR	NOUN

↓

I	NOUN
BOUGHT	VERB
JOHN	NOUN-POSS
CAR	NOUN

It will also handle sentences such as "Bill owns a jeep"-"George wrecked the jeep" but ignore sentences such as "Bill has a boat" -"Sam bought the car".

(b) The Conceptual Dependency Scheme of Schank (1969) involves mapping sentences into a language-free conceptual-network representation. The "conceptual parsing" of a sentence involves the elimination of "function" words as opposed to "concept-realizing" words, and the assigning of the concepts (feature-clusters?) of which the words are realizations to governing or subsidiary positions in a conceptual hierarchy. The relations between the concepts are denoted by graphic symbols indicating various kinds of two-way (equivalance) or one-way (governor-governed) linkages. To translate the LANAL-I notation into Shank's notation, it is necessary to realize that a link between two words can be represented by features of the words, either the same feature in both words (for a two way link) or different features (for the two ends of a one-way link). For example in "The boy hit the ball", "boy" and "ball" are "picture producers" (PPRD), "hit" is an "act" (ACTX), and the concepts realized by "boy" and "hit" are connected by a two-way link, indicated as boy <----> hit or PPRD <----> ACTX. This can be represented as:

BOY	PPRD-TWLK
HIT	ACTX-TWLK

in the LANAL-1 notation, where TWLK is a feature indicating a two-way linkage. The parsing rule involved here is simply:

(1)-PPRD + (2) - ACTX + X1 = (1) - TWLK + (2) - TWLK + X1.

Programming Details

LANAL-1 is written in the IBM-360 assembler language. The 360 is particularly useful for a program of this type, since it has a large repertoire of machine-instructions for logical processing.

The binary-number notation for the definition of a word, which was introduced above, is actually used as the internal representation of the word in the program. This makes context recognition, and adding or deletion of features, extremely simple to carry out by elementary Boolean Algebra operations.

If we are operating in terms of a ten-feature classification system, in which feature seven is NOUN and feature nine is ANIM (animate), looking for an animate noun as part of a context requires us to accept 0010001010 or 1100011110 but reject 1000010011. This can be done in terms of two ten-digit binary "masks".

X = 0000001010
Y = 0000001010

If we let Z represent the definition we are examining, the condition for acceptance is

Z and X = Y

where and represents the Boolean or logical product operator. If we were looking for the presence of feature seven and the absence of feature nine, the masks are

X = 0000001010
Y = 0000001000

Assuming that we have recognized an animate noun as part of a context, let us mark it positively for feature 4, hich is SUBJ (subject). Letting Z represent the original definition and W the transformed definition, the operation is simply:

W = Z or X

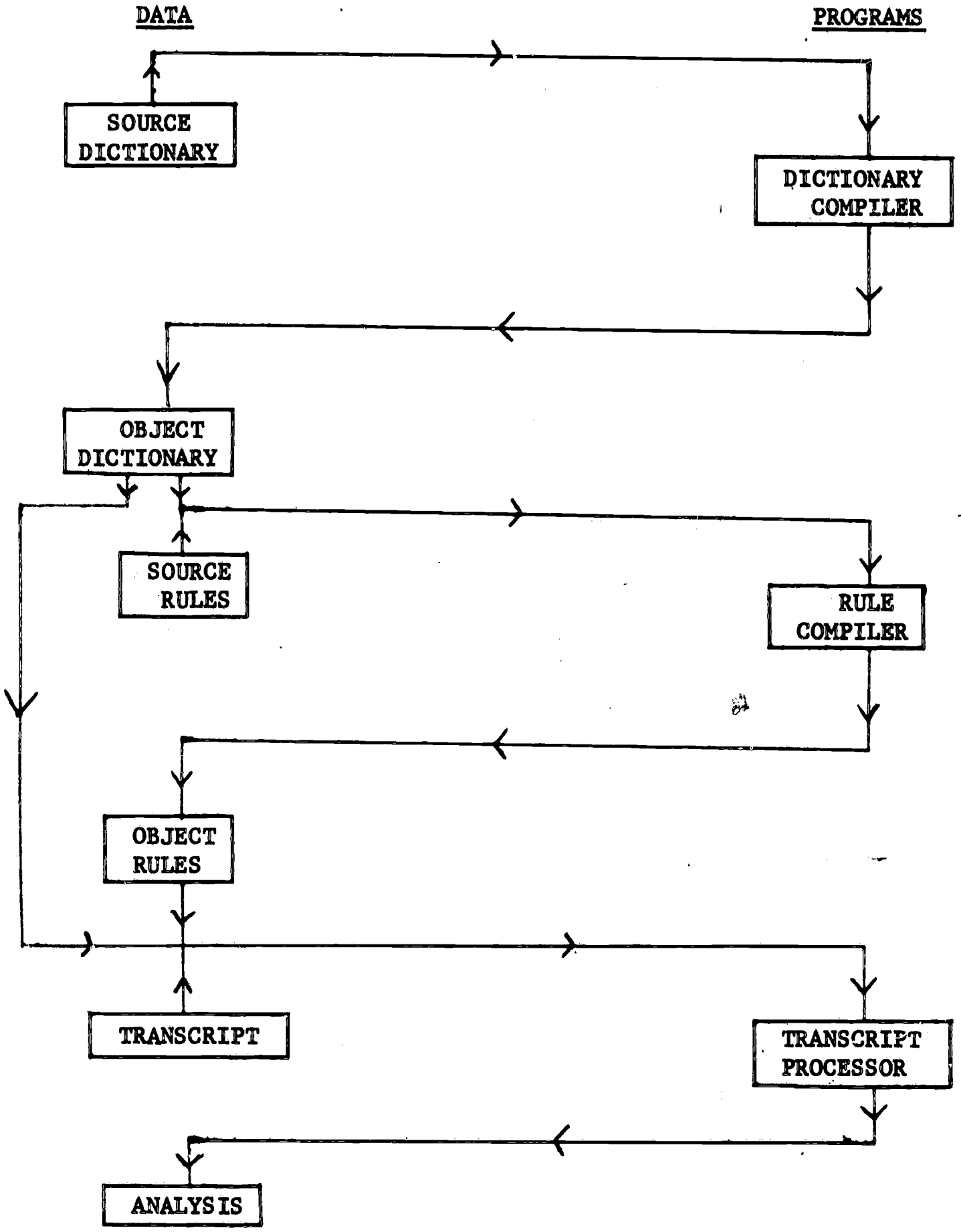
where X is the mask 0001000000, and or the logical sum operator.

Structure of the Program

The overall structure of the program is illustrated in Figure 2. It is actually composed of a two-part compiler (the dictionary compiler and the rule compiler) and an operating system (the transcript processor). The source dictionary is input to the dictionary compiler, which condenses it considerably, performs certain referencing operations to speed up word look-up, and produces an "object dictionary." This "object dictionary", together with the "source rules" (that is, the rules written in the rule shorthand), are input to the rule compiler. Under the control of the object dictionary, the rule compiler translates the rule-shorthand symbols into the binary masks used in the actual transcript processing, and outputs an "object rule" set. The transcript processor accepts as input the object dictionary, the object rules, and a transcript, and uses the dictionary and rules to operate on the transcript and produce the "analysis" or output, which is in the sentence notation used above. Because of the multistep design of the program, considerable time and space may be saved on certain jobs. For example, if the same dictionary and rule system are to be used in a number of analyses, the source dictionary and source rules need not be recompiled, and the transcript processor may run by itself, using the previously compiled object rules and object dictionary as input. Similarly if the rules, but not the dictionary, are to be changed, the dictionary compiler need not be rerun.

Figure 2

Structure of the LANAL-1 Program



Appendix C

Coding of Verbalizations

Punch Card Format

Data were transferred to the punch cards in the following form:

Column Identification data	Punch Code		Description
1	1		Washington Square Sample
	2		West Harlem Sample
	3		Cross-Sectional Sample
2	1		Age 2-1/2 or session #1
	2		Age 3 or Session #2
	3		Age 3-1/2 or session #3
	4		Age 4 or session #4
	5		Age 4-1/2 or session #5
	6		Age 5 or session #6
3, 4	01		Mother-child pair #1
	.		
	.		
	12		Mother-child pair #12
Utterance Number 5-8	0001		Utterance #1
	.		
	.		
	0999		Utterance #999
Speaker of utterance 9	1	M	Utterance by mother
	2	C	Utterance by child
An utterance was identified as codable or uncodable:			
10	1		Utterance can be coded, continue and code col. 11-16
	2	OMIT	Utterance cannot be coded. Contains blanks in transcription. (Give me the _____)
	3	OY	Utterance cannot be coded. Is incomplete, incomprehensible. (crying, laughing, noises....)
	4	OC	Utterance cannot be coded. It is a repetition of a statement just for the sound of it.

	5	(This is a seat, seat, seat, seat) Gap. When a behavioral act is responded to, it is necessary to put in an extra line so that the response can be recorded.

All utterances which could be coded (1 in col. 10) were then coded for the following characteristics: Form, Content of information, Mode of Information Exchange.

Column No.	Punch Code	Alphabetic	Description
		Form	
11	1	Q	Question
	2	S	Statement
12 (if 1 in col.11)	1	Q	Questions which request information. All "wh" questions (What is this?)
	2	H	Questions which present a hypothesis, and can be answered by yes or no. (Is this a blue car?)
	3	A	Questions which request attention (See this car?)
	4	B	Questions which request behavior. (Will you get it for me? Shall we play with the car?)
(if 2 in Col.11)	1	S	Statements which provide information. (This is a magnet board.)
	2	H	Statements which present a hypothesis. Speaker is usually uncertain, doubtful. (I think this is it.)
	3	A	Statements which seek to elicit attention. (Look at this. Watch.)
	4	B	Statements which seek to elicit behavior. (Get that piece of track. Let's go home.)

Information Content

13	1	P	Permanent information is given about rules, labels, or characteristics which hold true in other situations and at other times. This category has highest priority when a double code might be possible. (The car is blue. The window must be kept closed all the time.)
	2	T	Transient information: utterances containing information which is true or relevant only in the present situation at a given time. (The lady will be right back. I want the car. The window is open. You don't know it?)
	3	B	Behavioral information:remarks which carry no other information than comments on the actions of the people involved or the person's own actions. (Come over here. Get off the table. Let's go.)
	4	F	Fantasy information:remarks about pretend situations. (They [the doll family] are going on a picnic.)
	5	A	Approval: approval is expressed without other information being conveyed. (You are doing really well. I like the way you fixed it.)
	6	D	Disapproval (I don't like you to do that. Naughty boy.No, no don't do it.)

Mode of Information Exchange

14-15	01	a	Explains: Gives or requests information about causality, function, purpose, goal or intention. (Why is it raining? You can't do it because it's dangerous. What is that button for?)
	02	b	Limits: Gives or requests information about the possibilities or constraints inherent in a situation. (You may not scream here. This is too large for the space. I'm not big enough to do this yet. There are no balls here. The track can't go there.)

Col. No.	Punch Code	Alphabetic Code	Description
	03	c	Clarifies: Gives or requests further information or repetition. (What did you say? This is the one I mean.)
	04	d	Describes: Statements or requests which deal with physical appearance, description, representations of actions. (Cats' eyes shine in the dark. Is this bigger than that? I'm going to pick them up.)
	05	f	Feeling: Statements or questions which describe feelings, wants. (I want to go home. I am tired.)
	06	g	Demonstrates: Shows how to do something while talking about it. (It goes into the puzzle this way. Show me how to do it.)
	07	k	Commands: Requests with insistence, anger, emphatically orders. Usually has exclamation marks. (Get right down now! You must do it immediately. You come here! Come on!)
	08	l	Labels: Gives or requests the name of an object or person. No other information given except the name. (This is a dog. What is this?)
	09	s	Specifies: Requests or provides information about location and selection. Usually can be conceived of as accompanied by pointing. (I will get this one. Go over there. Which one shall I do.)
	10	t	Thematic: Sounds made as part of the "script" in fantasy (choo choo, Mooo.)
	11	r	Reiteration: Repetition by same speaker, provides no new information. Not to be confused with "03" which is used when the responding person has asked for repetition. Need not be consecutive to the previous utterance. (M: What is that? M: What is it, Anthony?)
16	1	0	Not end of topic.
	2	Z	End of topic. Code next to the last utterance of the previous topic. Change of topic is generally when referent changes. i.e., puzzles to lock box, train to balance beam.

Col. No.	Punch Code	Alphabetic Code	Description
Utterances were next coded as Response Demanding or not			
17	1	+	Response demanded by this utterance. (Will you put that back? Go to the cabinet.)
	2	-	Response not demanded by this utterance. (I am taking the cars out now. I have two trains.)
Utterance Responded to or not			
18	1		Utterance was clearly responded to. (M: Do you want to play with this? C: Yes)
	2	00	Utterance was ignored. This is usually ascertained by consulting the behavior language. (M: Do you like the doll? C: [no answer])
	3	OY	Response to utterance was unclear. (M: Can you do this? C: Umm)
	4	OM	Utterance following is coded 2 in Col. 10 (OMIT) so it cannot be clearly coded for quality.
	5		Speaker continues to speak or exchange continues. (a) other person does not have time to answer (b) a response is not required. (M: Do you like the airplane? M: Hmm? [said before child has time to answer])
	6		The first utterance in a new exchange. An exchange is a series of stimuli and responses about one thing. M: You spell my name with a "I". C: (sound) C: Let's make... M: Hmmm? C: Spell your name. M: Come on. M: What ya looking at? (Code 6) C: Watching burn up place. new exchange begins).

Col. No.	Punch Code	Alphabetic Code	Description
			Quality of Response
			Code only if 1 was coded in Col. 18, i.e. utterance was responded to clearly.
19	1		Direct information was present in response. (M: Do you like this? C: Yes)
	2		Peripheral information was present. (What's that? Ask your father.)
	3		No information (That's a dog? What did you say?)
20-21			Code if 1 (direct) or 2 (peripheral) is coded in column 19.
	10		Misinforms Gives incorrect information. (It's a horse [when it is really a calf]).
	01		Informs: Gives factual information as requested. (What is that? It is a ball.)
	02		Confirms: Simple agreement in response to a question or statement. (This is green. Yes, it is.)
	03		Denies: Declares a remark to be untrue without correction. (Today is Tuesday. No, it's not.)
	04		Corrects: Provides correct information. (This is blue. No, it's purple.)
	05		Accepts/complies: Accedes to a request or accepts information or opinion. (Okay, I will do it. That's alright with me.)
	06		Rejects: Non-acceptance or non-compliance (No, I won't help. I don't want that one.)
	07		Praises: (That's beautiful. You did it just right.)
	08		Punishes: Indicates disapproval or dissatisfaction. (That's bad. Very naughty.)
	09		Uncertainty: Indicates doubt about information provided or in general. (I'm not sure. I think it might be true.)

Col. No.	Punch Code	Alphabetic Code	Description
22	Code if 3 (no information) was coded in column 19, leave columns 20-21 blank		
	1		Direct question in response to the preceding utterance. (What did you say? What do you mean? I do what?)
	2		Peripheral question in response. (This is a balloon. Where did you buy it?)
	3		Remark reflected back. (What do you think of it? You tell me.)
	4		Encouragement. (Try again. Keep on pushing. Go ahead.)
	5		Parroting. (M: There it is. C: There it is.)
	6		"Don't know" response.
	7		Delayed. (I'll tell you later) or answers self immediately.
	8		Repetition in response to an ignoral. (M: What is that? C: [ignores] M: What is it Anthony?)
Utterance Length			
23-24	00 . . nn		Number of words in each sentence.
All utterances were coded with respect to aspects of the playroom that were being TALKED ABOUT.			
25-26	00		Not coded
	01	U	Uncodable
	02	Beh	Behavior -- centered upon behavior
	03	Sit	Situation -- centered on situation outside playroom
	04	BB	Balance Beam
	05	B	Barn
	06	H	House
	07	PA	Airplane Puzzle
	08	PS	Ship Puzzle
	09	PW	Street Puzzle
	10	P	Puzzle (unspecified)
	11	LB	Lock Box
	12	T	Train
	13	ML	Magnetic Letters
	14	MC	Magnetic Cars
	15		Double Coded -- Two or more objects were being talked about or played with simultaneously.

Col. No.	Punch Code	Alphabetic Code	Description
Utterances were also coded with respect to what the subjects were playing with.			
27-28	00 . . 15		Same code as was used above with talking about.
Sentences were coded as either transformed or kernel sentences.			
29	0 1 2	 K T	Not coded Kernel sentence Transformed Sentence
If a sentence contained transformations, both the number of single and double based transformations were counted.			
30-31	00 . . nn		Number of Single Base Transformations
32-33	00 . . nn		Number of Double Base Transformations
Sentence complexity was specified by summing the weighted units assigned to each transformation or kernel.			
34-35	00 . . nn		
Columns 36-71 are divided into 18 two-column fields. Transformation code numbers are listed across the card, one transformation to a two-column field. A maximum of 18 transformations may be listed for any one particular sentence. See page , appendix for a list of 89 transformations with code numbers, assigned complexity weights, and examples.			
36-71	00 . . 89		

Col. No.	Punch Code	Alphabetic Code	Description
72-74	BLANK		

Socio-Economic-Status was coded based on Hollingshead's two factor index.

		Social Class
75	1	I -- upper
	2	II
	3	III
	4	IV
	5	V -- lower

Sentence numbers before and after coding were not always the same because of numbering errors or changes made by the coder. Therefore, the original sentence number identifying the sentence on the typed transcript was transferred to the punched cards.

76-80	Original sentence number found on the typed transcript. Column 80 is an alpha field to account for new numbers inserted into the transcript.

Appendix D
Sample and Age Differences
by VINEX categories

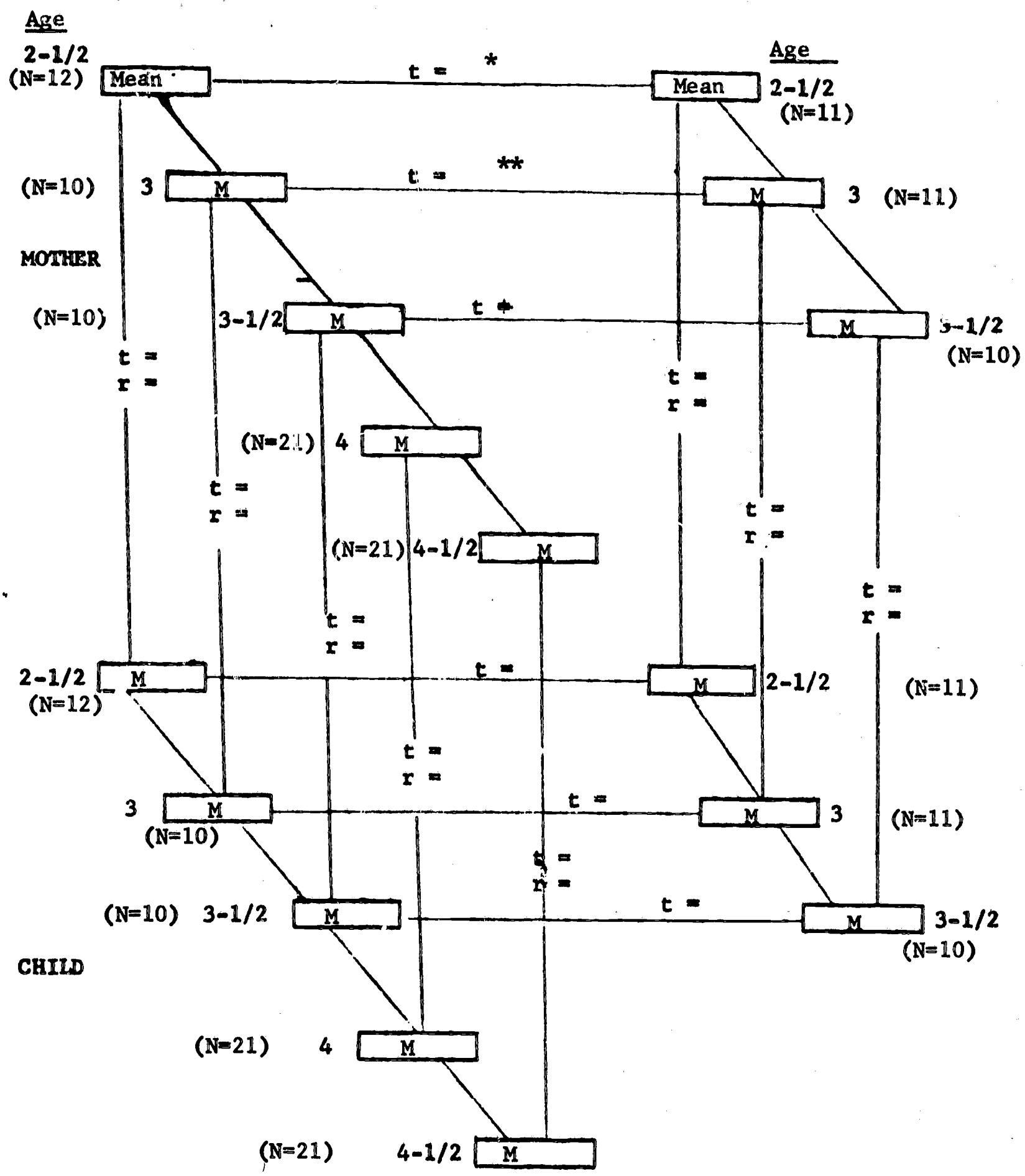
Tables 1-37 Means, correlations, t's and Significance: All groups, mothers and children on each VINEX category

Appendix D

Table 1

Variable: The following tables contain the Means, Correlations, t's and significance for all groups, mothers and children on each VINEX variable.

Harlem **Washington Square**



1. * Significance < .05
   ** Significance < .01
2. r's - correlation between mothers and children within samples

Table 2

Variable: Individual Utterance Frequency

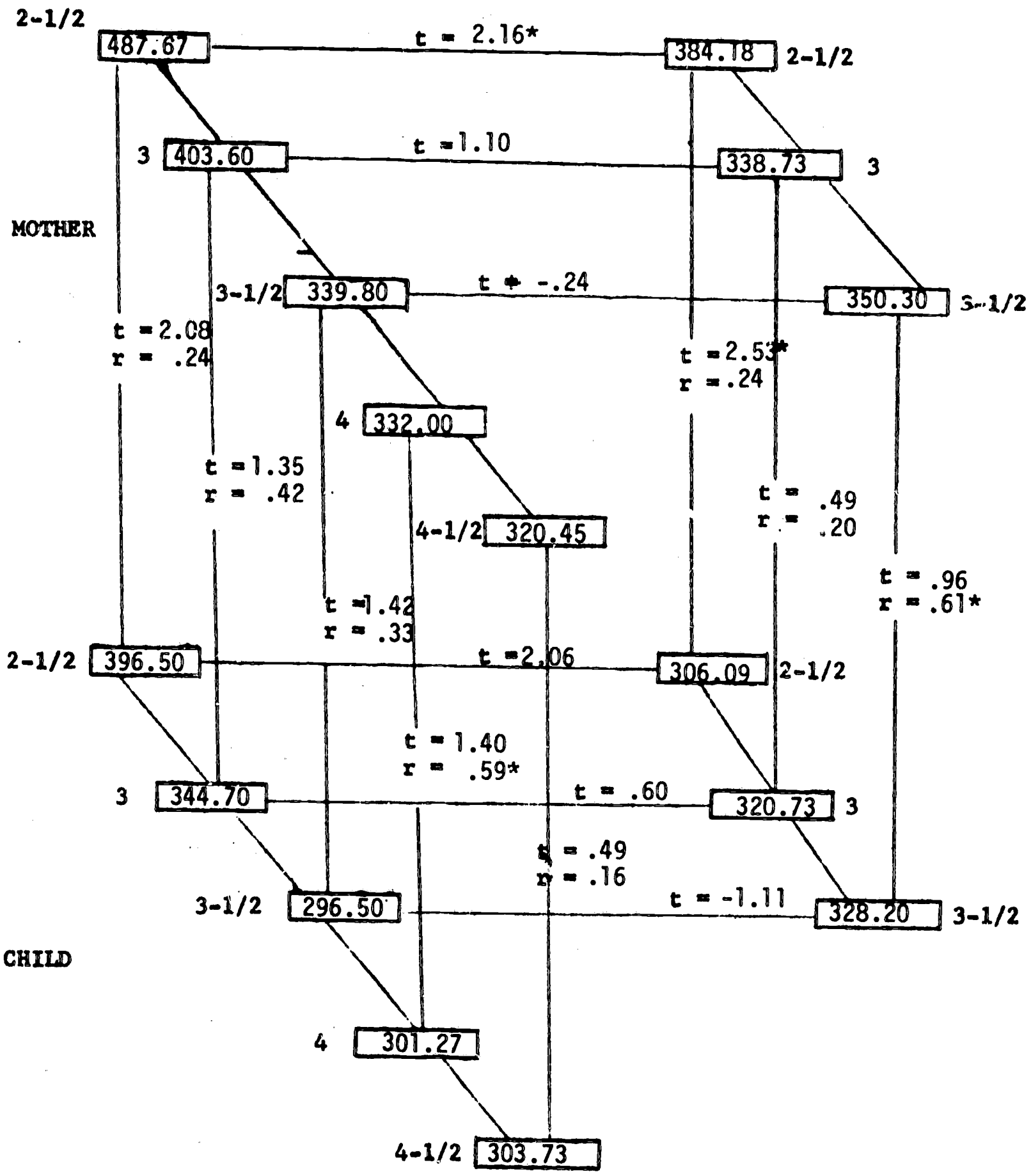


Table 3

Variable: Coded Utterance Frequency

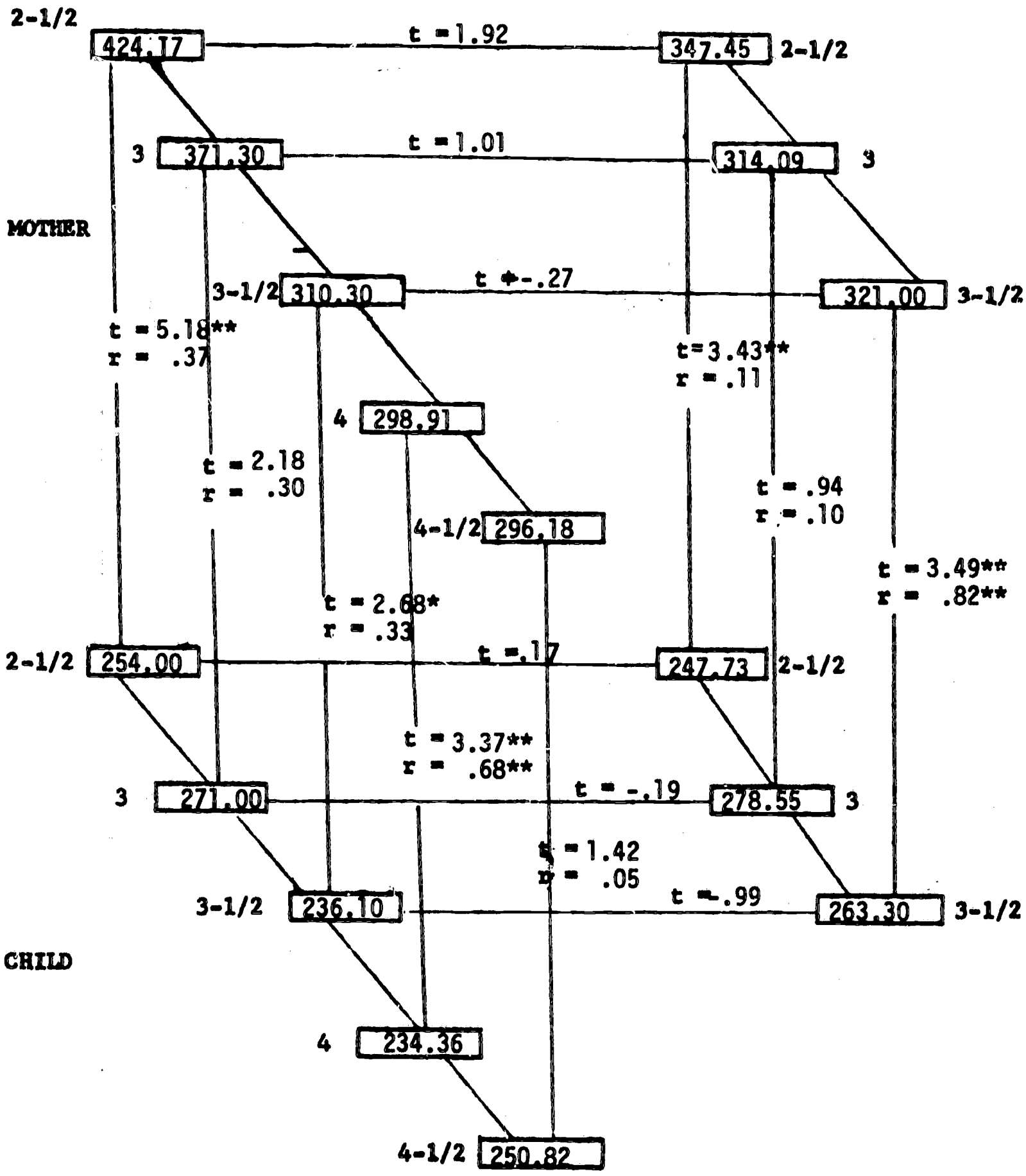


Table 4

Variable: Question Frequency

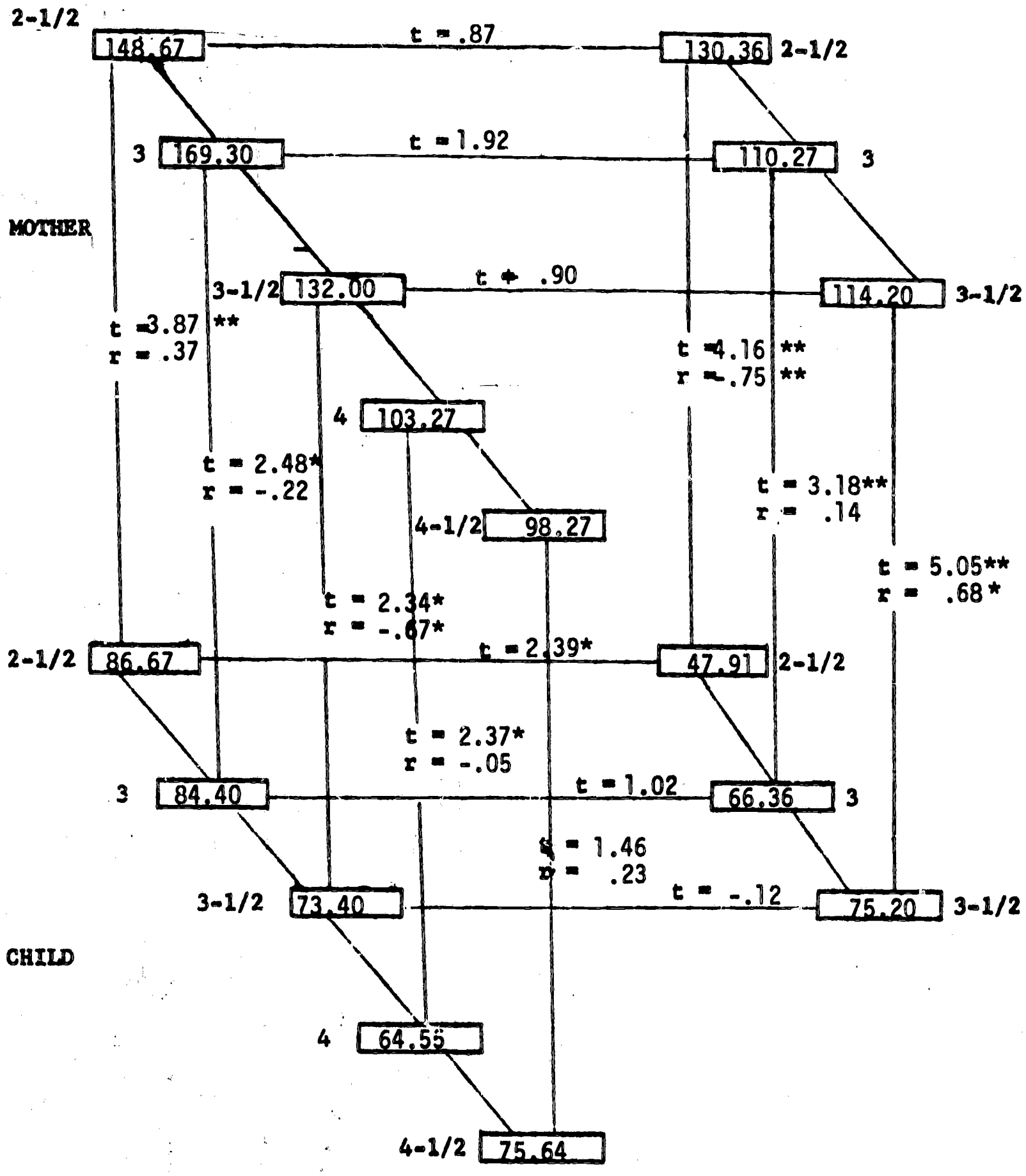


Table 5

Variable: Statement Frequency

Harlem **Washington Square**

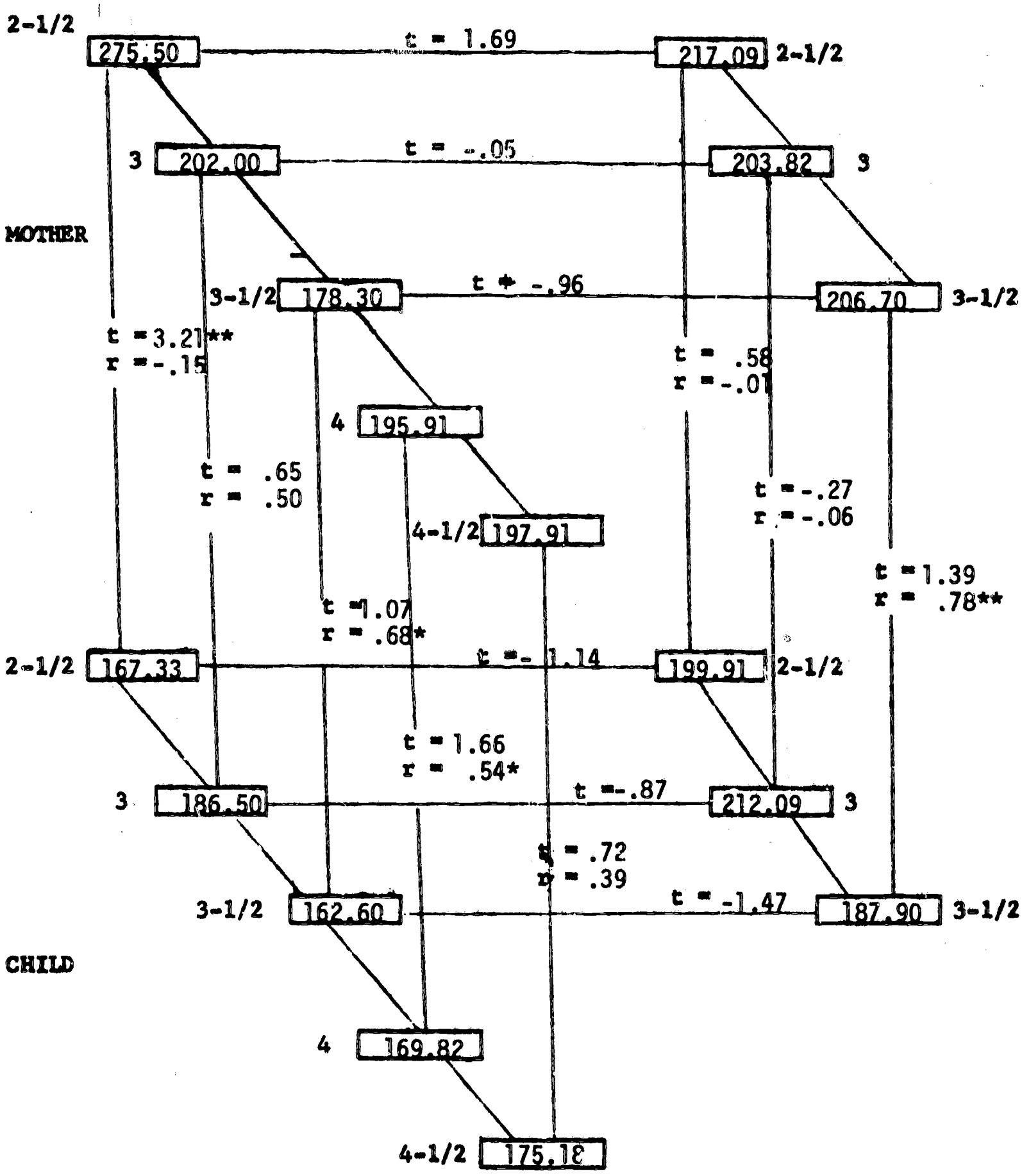


Table 6

Variable: % Utterances Coded

Harlem **Washington Square**

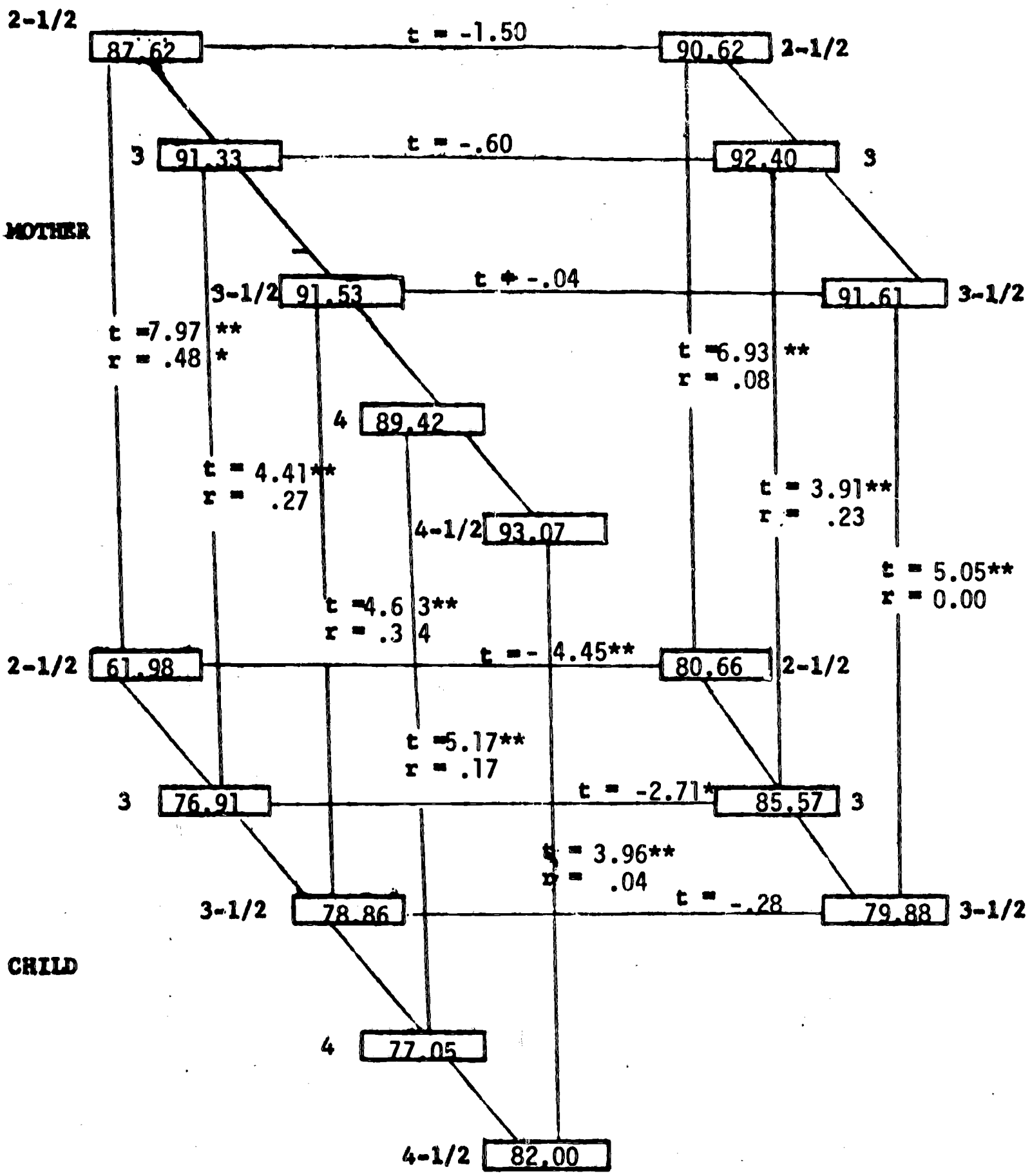


Table 7

Variable: Form- % Questions

Harlem **Washington Square**

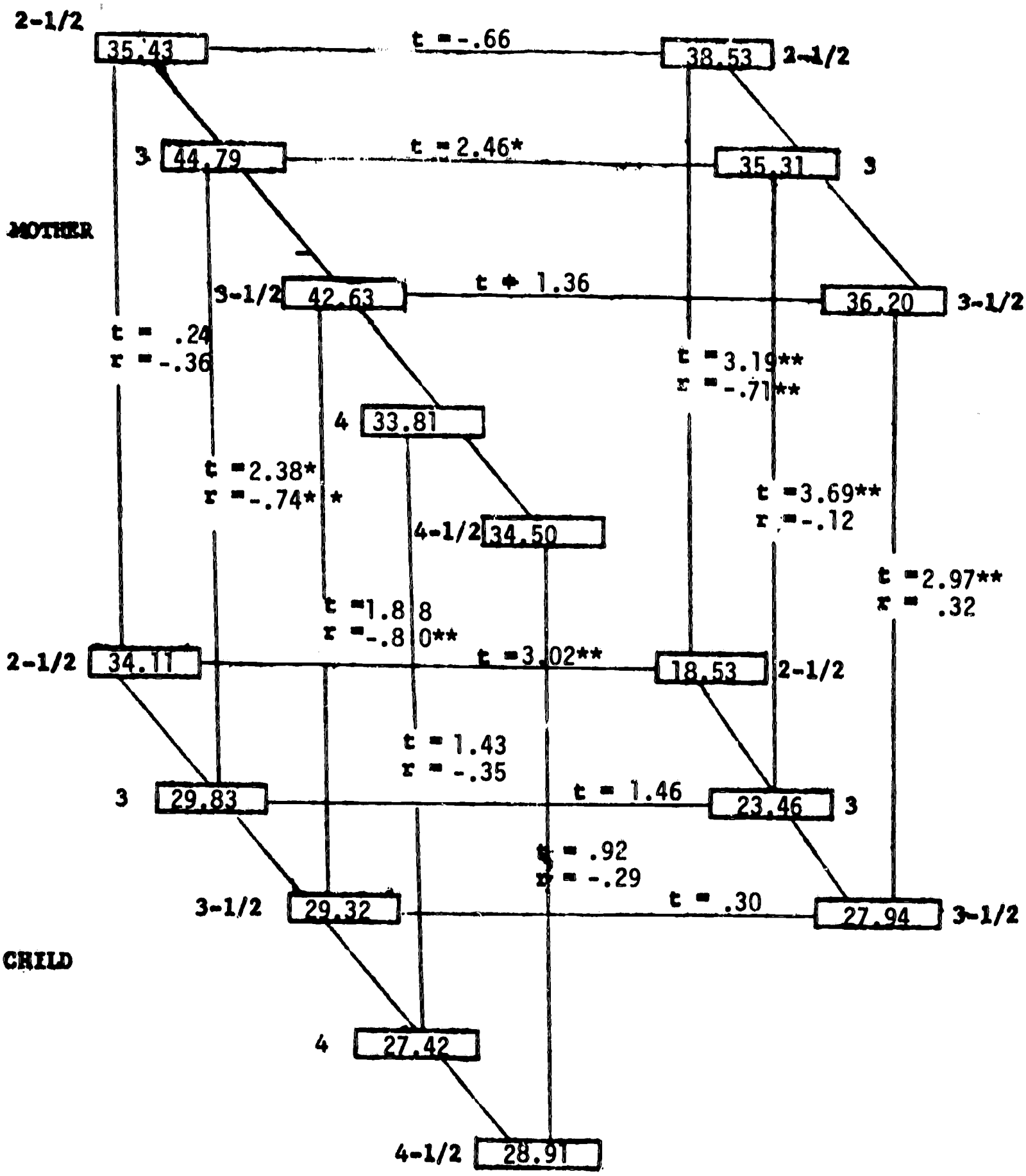


Table 8

Variable: Form - % Statements

Harlem **Washington Square**

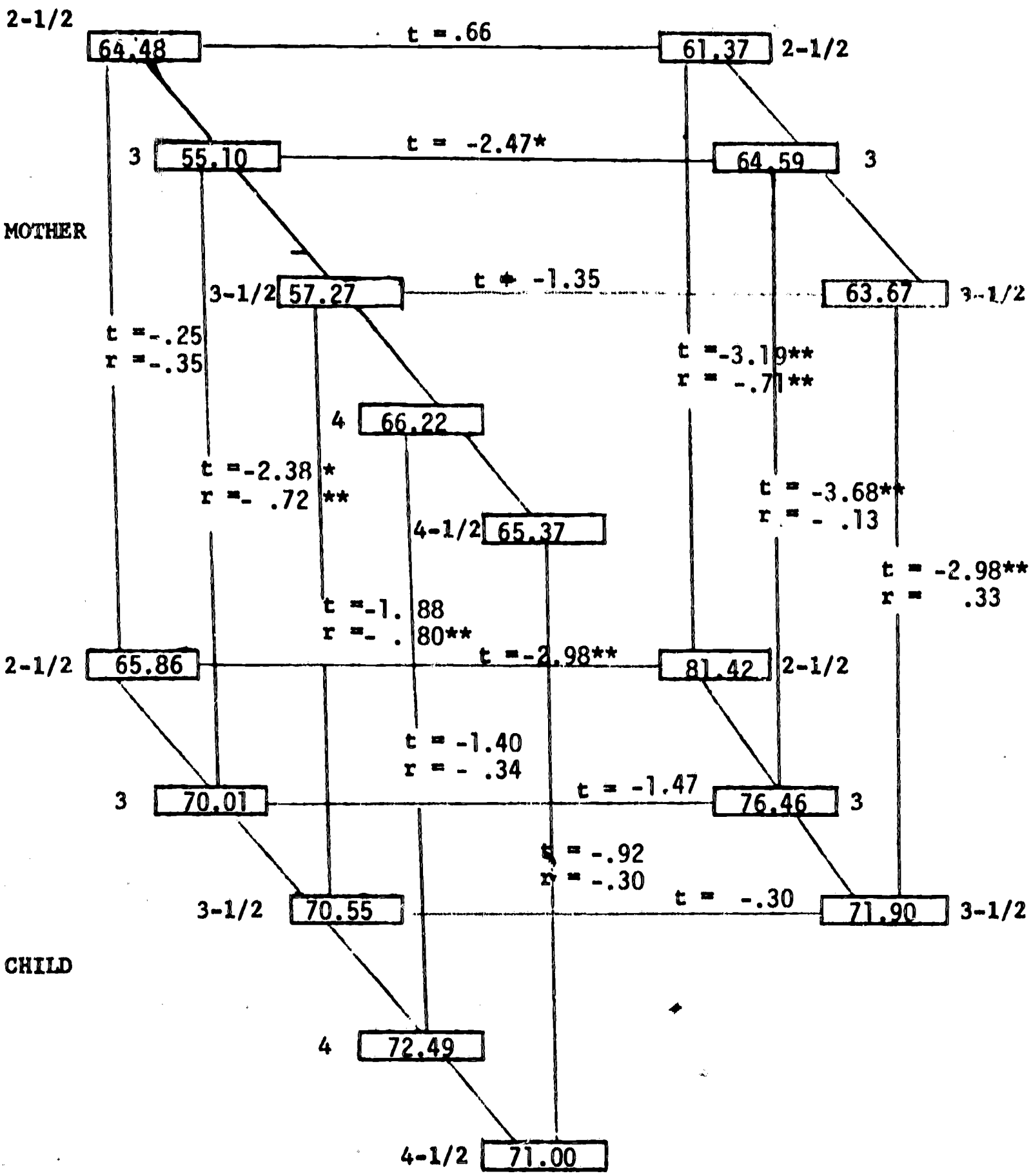


Table 9

Variable: Content - % Permanent Information

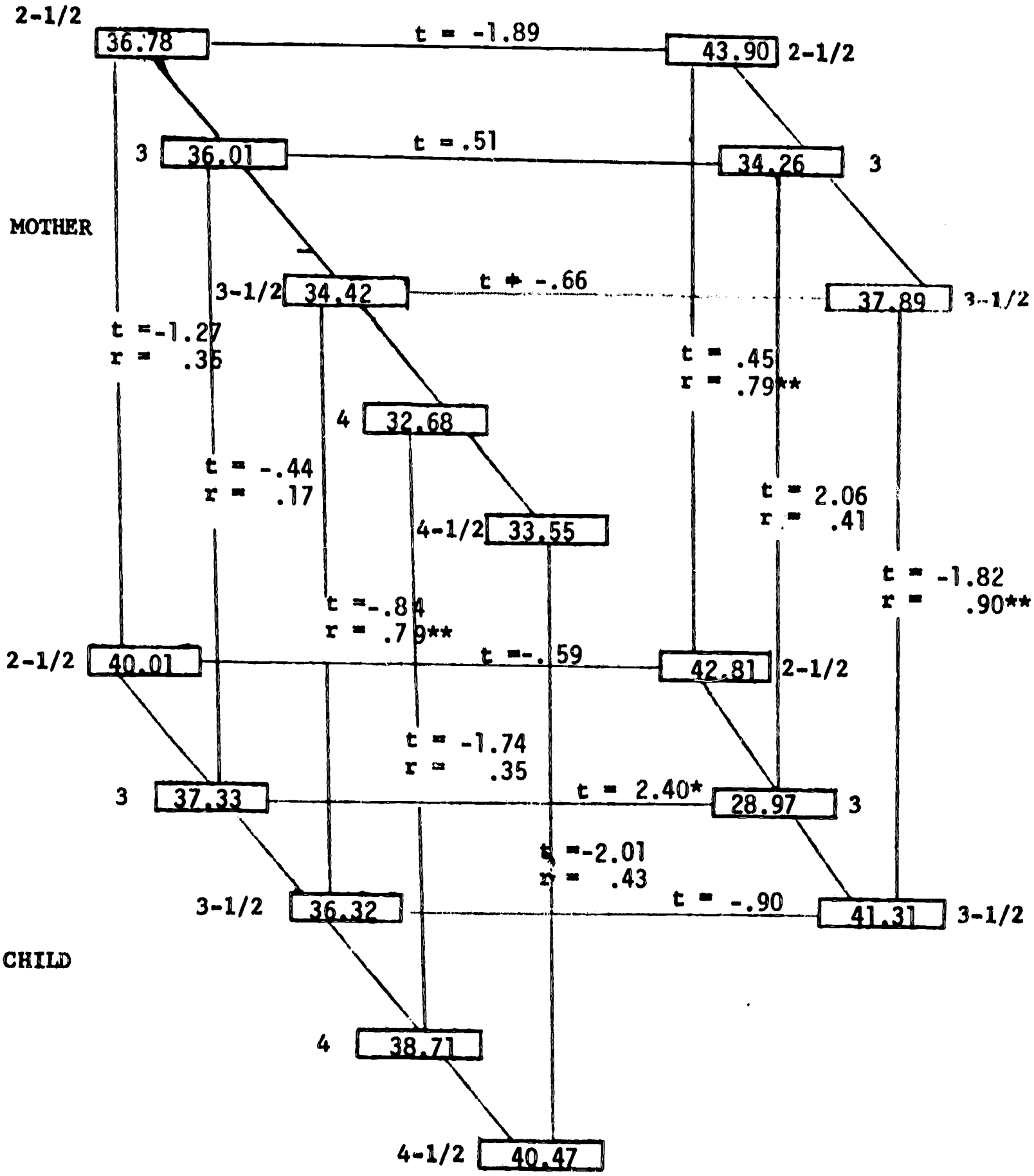


Table 10

Variable: Content - % Transient Information

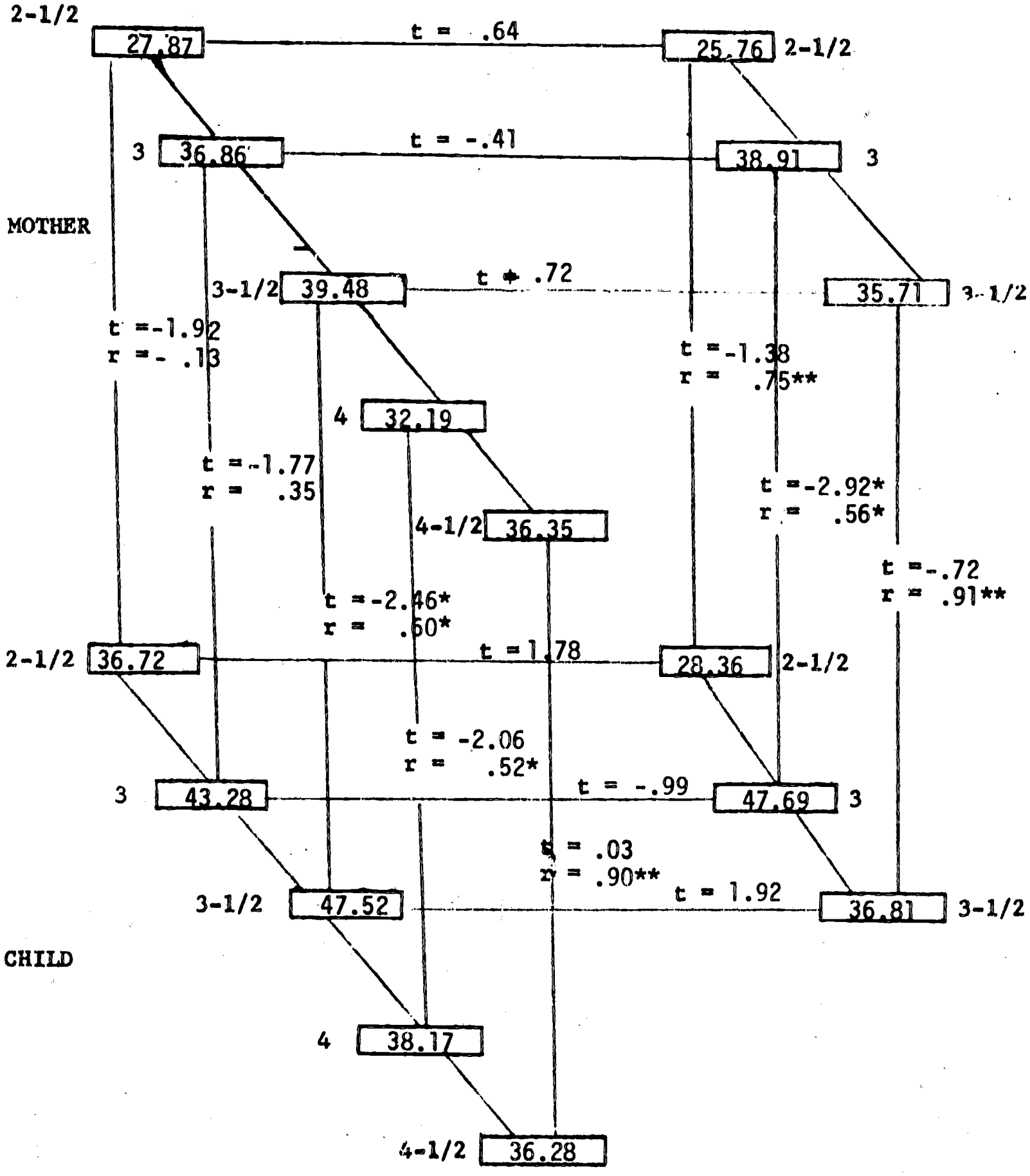


Table 11

Variable: Content - % Behavioral Information

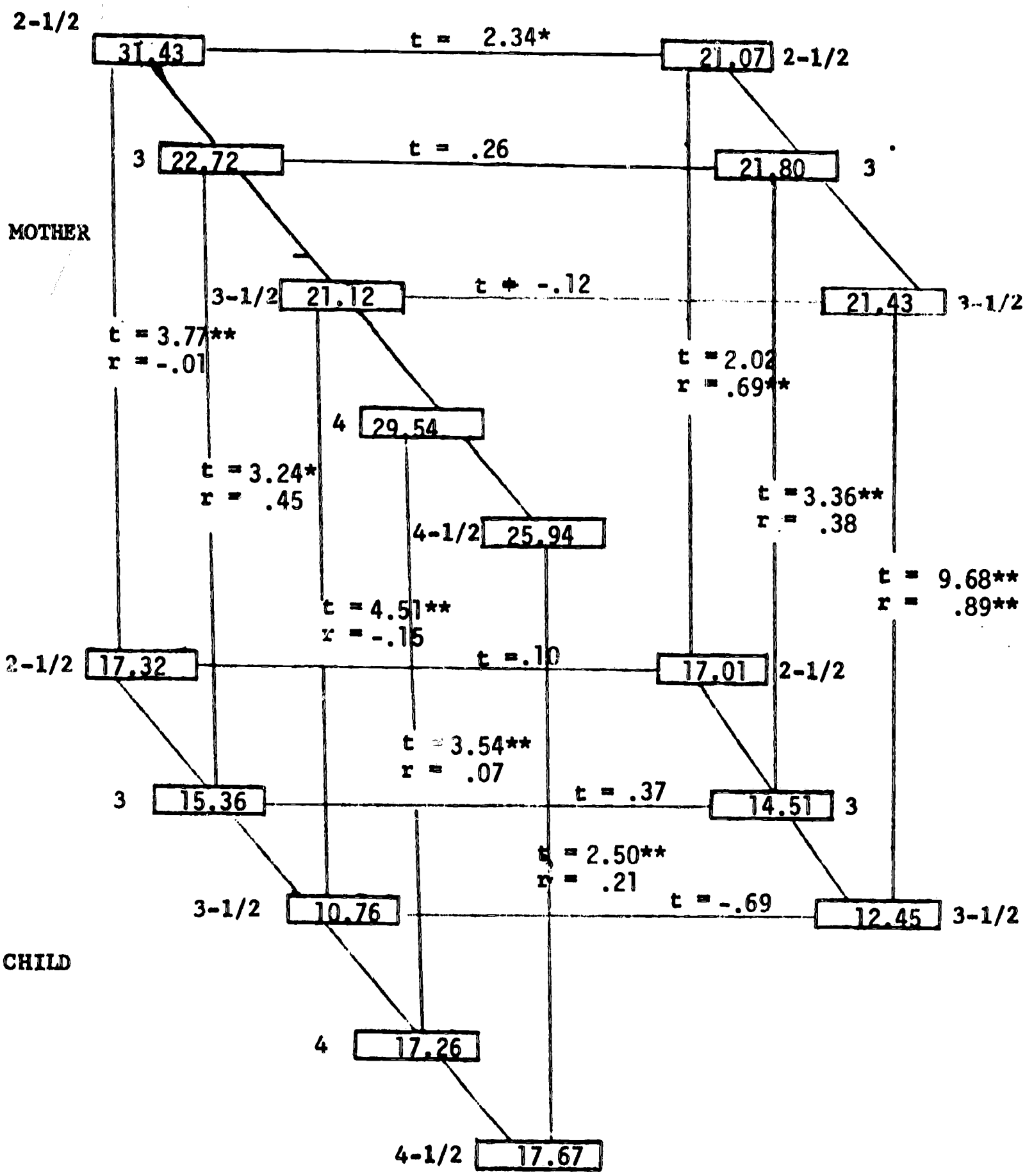


Table 12

Variable: Content - % Fantasy

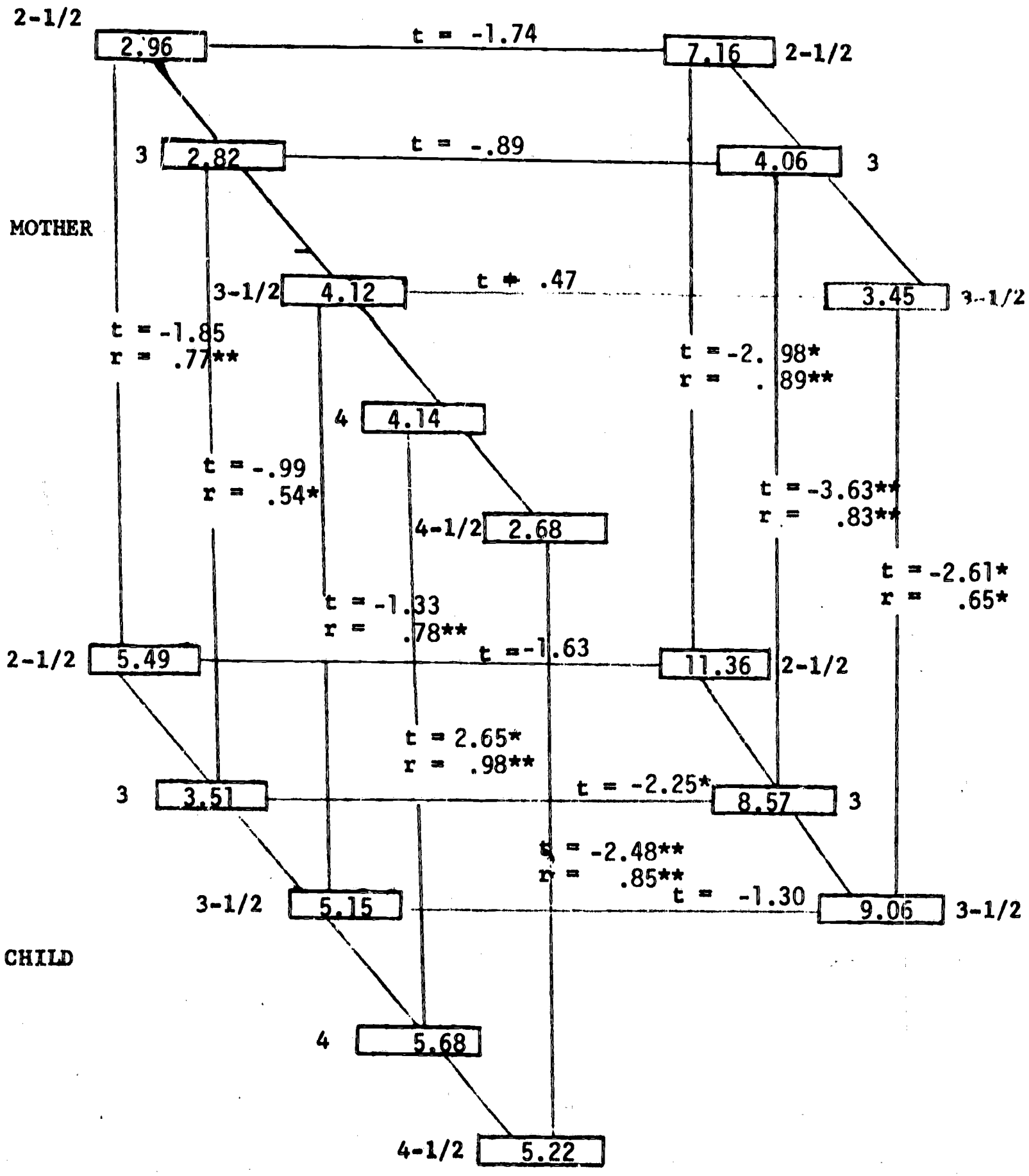


Table 13

Variable: Mode - % Explanations

Harlem **Washington Square**

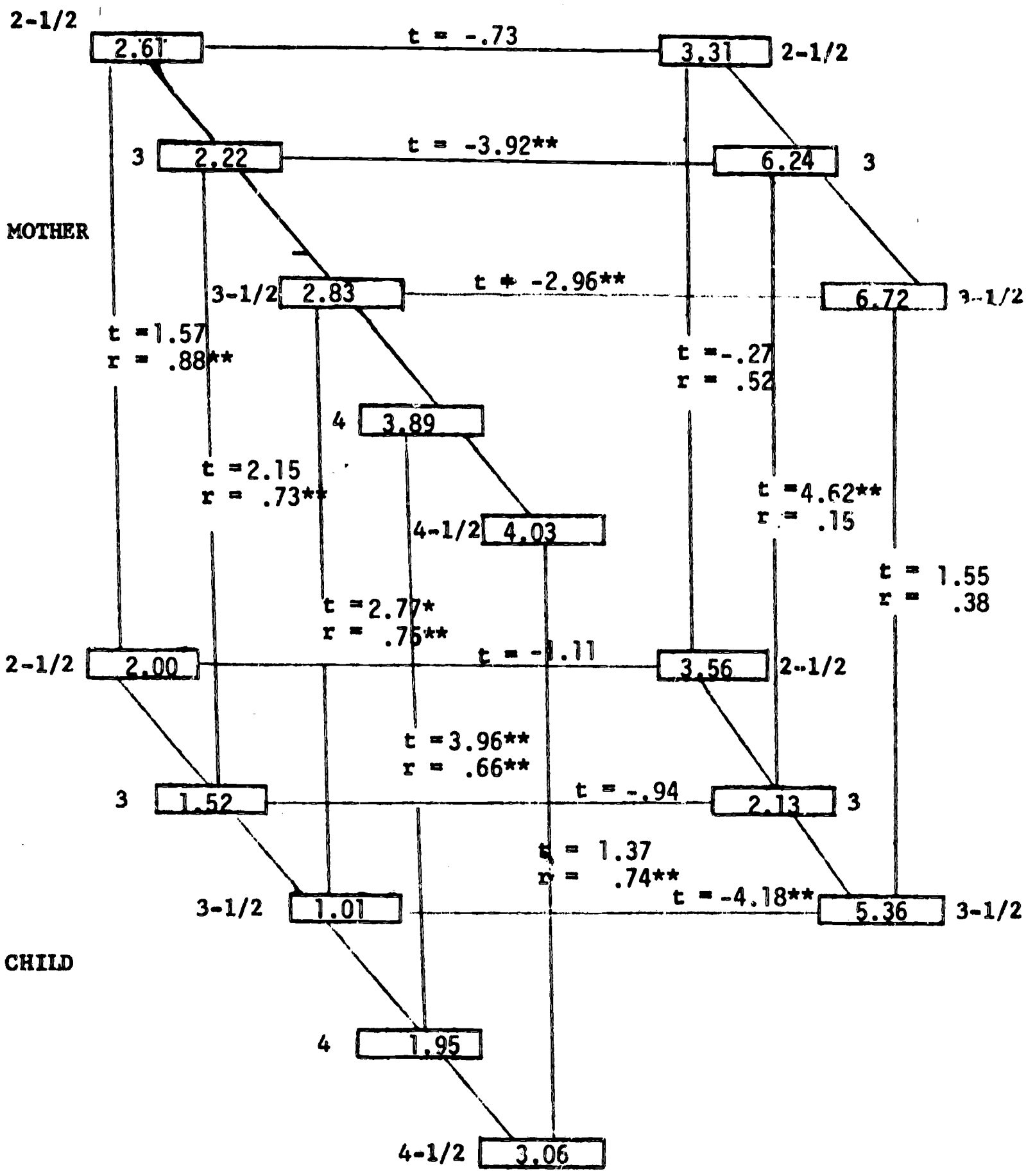


Table 14

Variable: Mode - % Limits

Harlem **Washington Square**

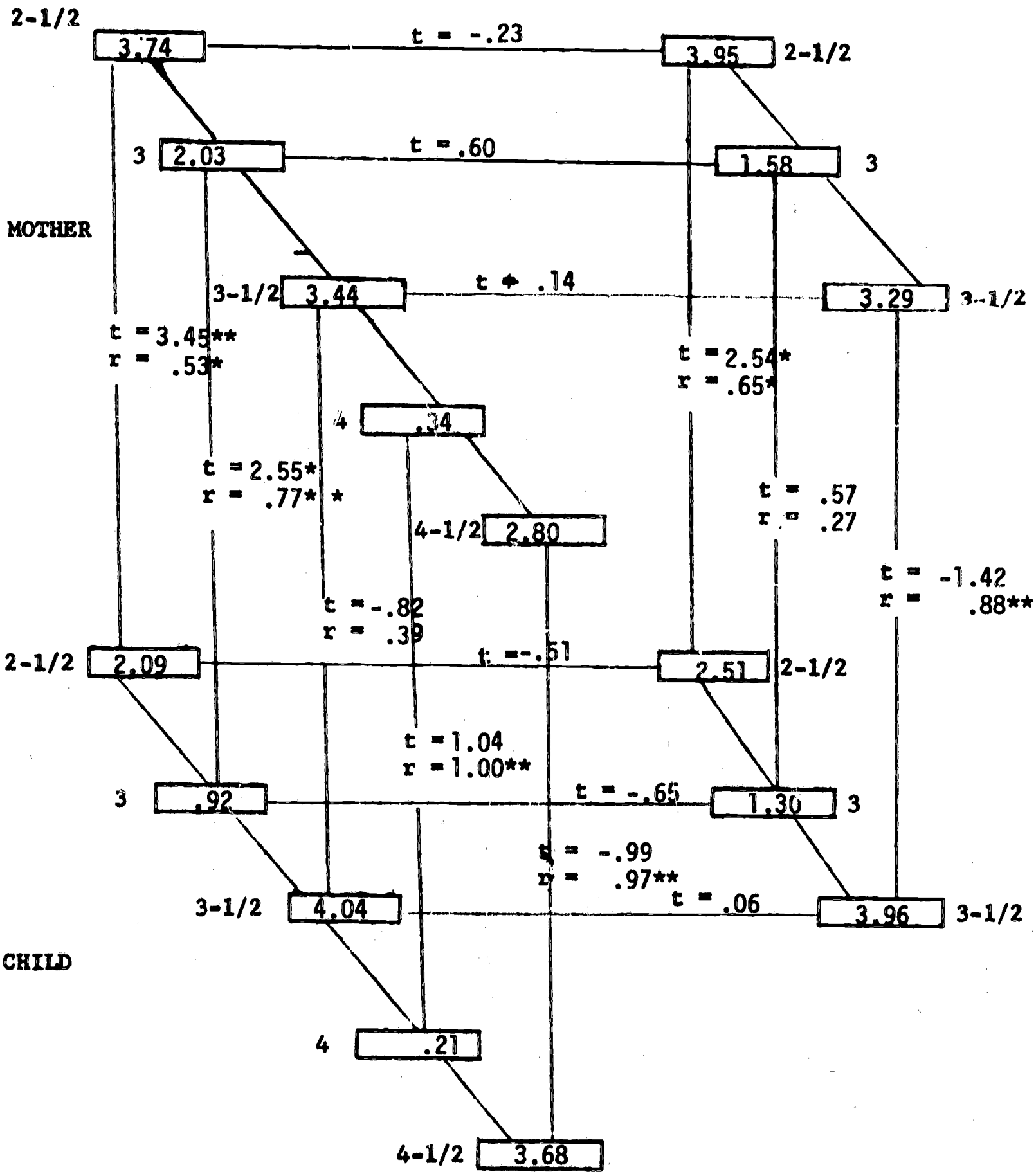


Table 15

Variable: Mode - % Clarifications

Harlem **Washington Square**

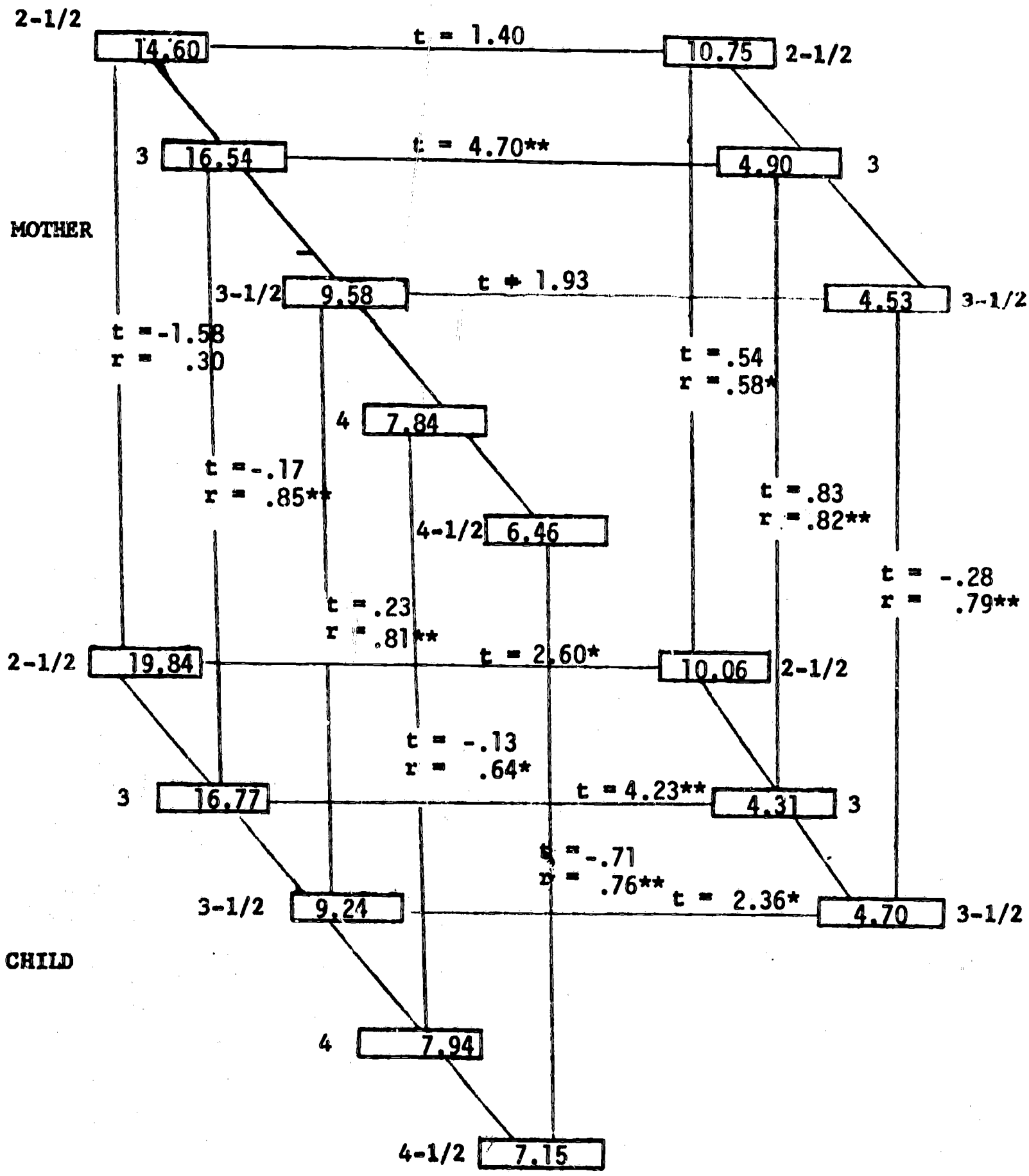


Table 16

Variable: Mode - % Description

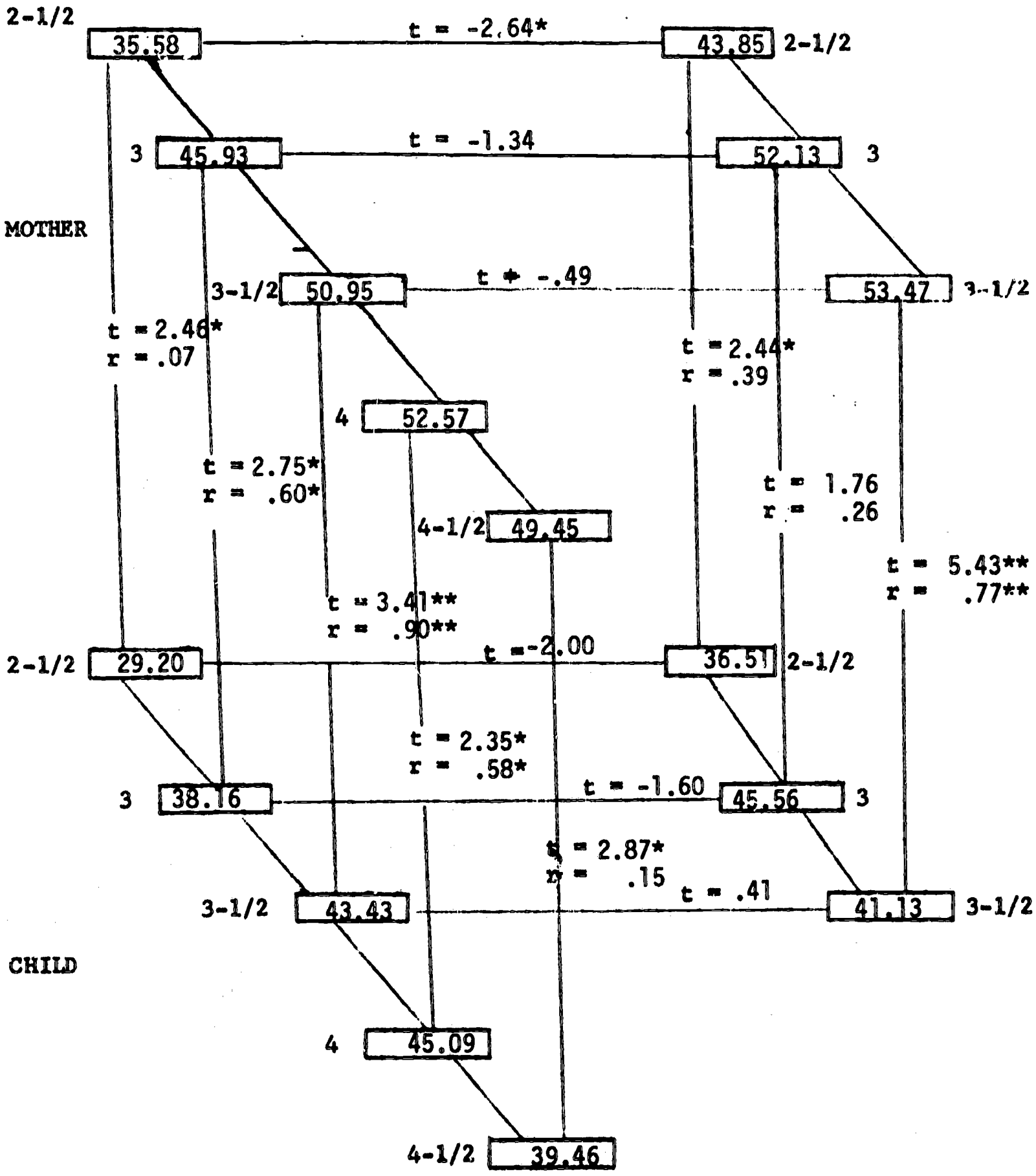


Table 17

Variable: Mode - % Feeling

Harlem **Washington Square**

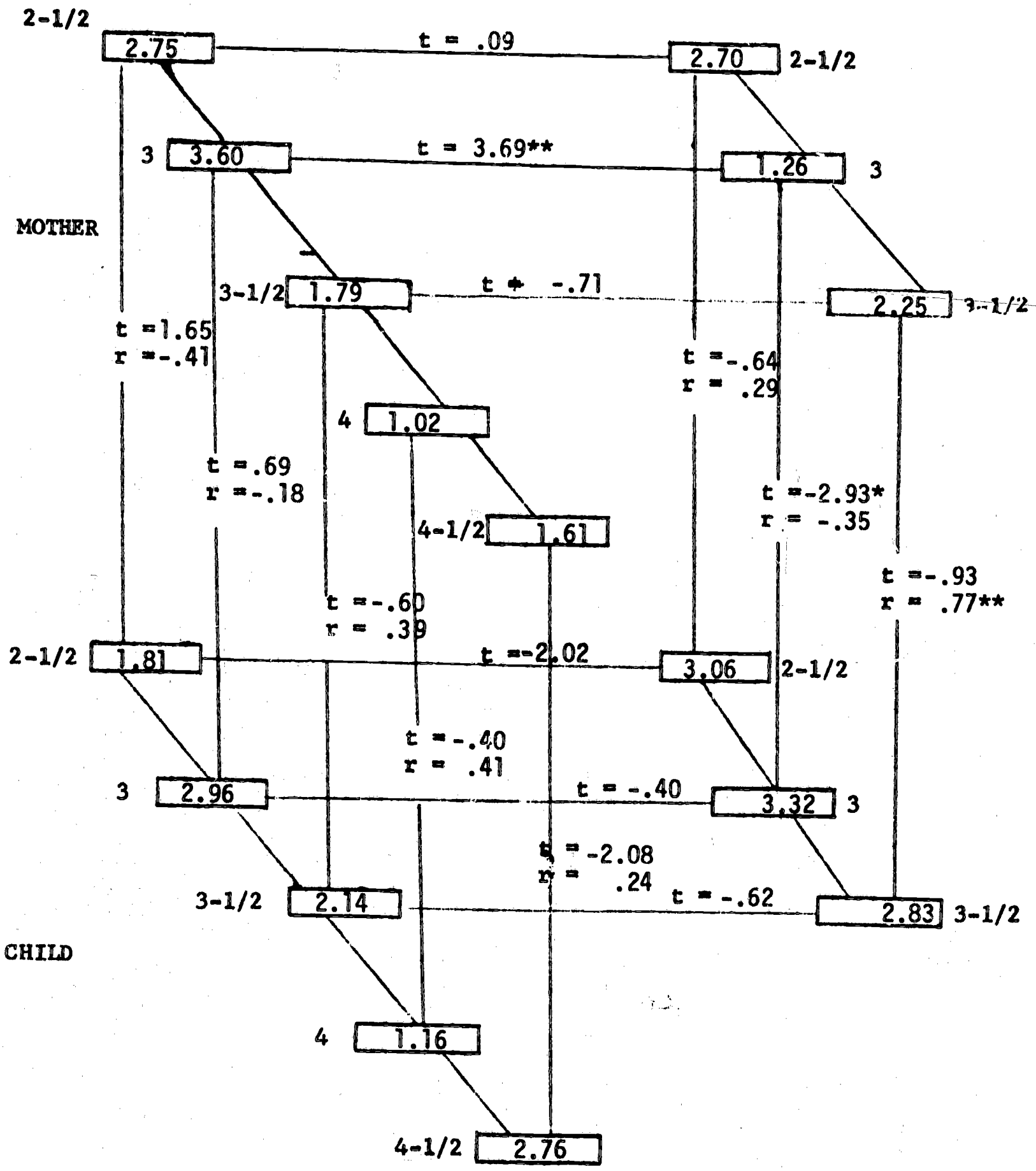


Table 18

Variable: Mode - % Demonstrates

Harlem **Washington Square**

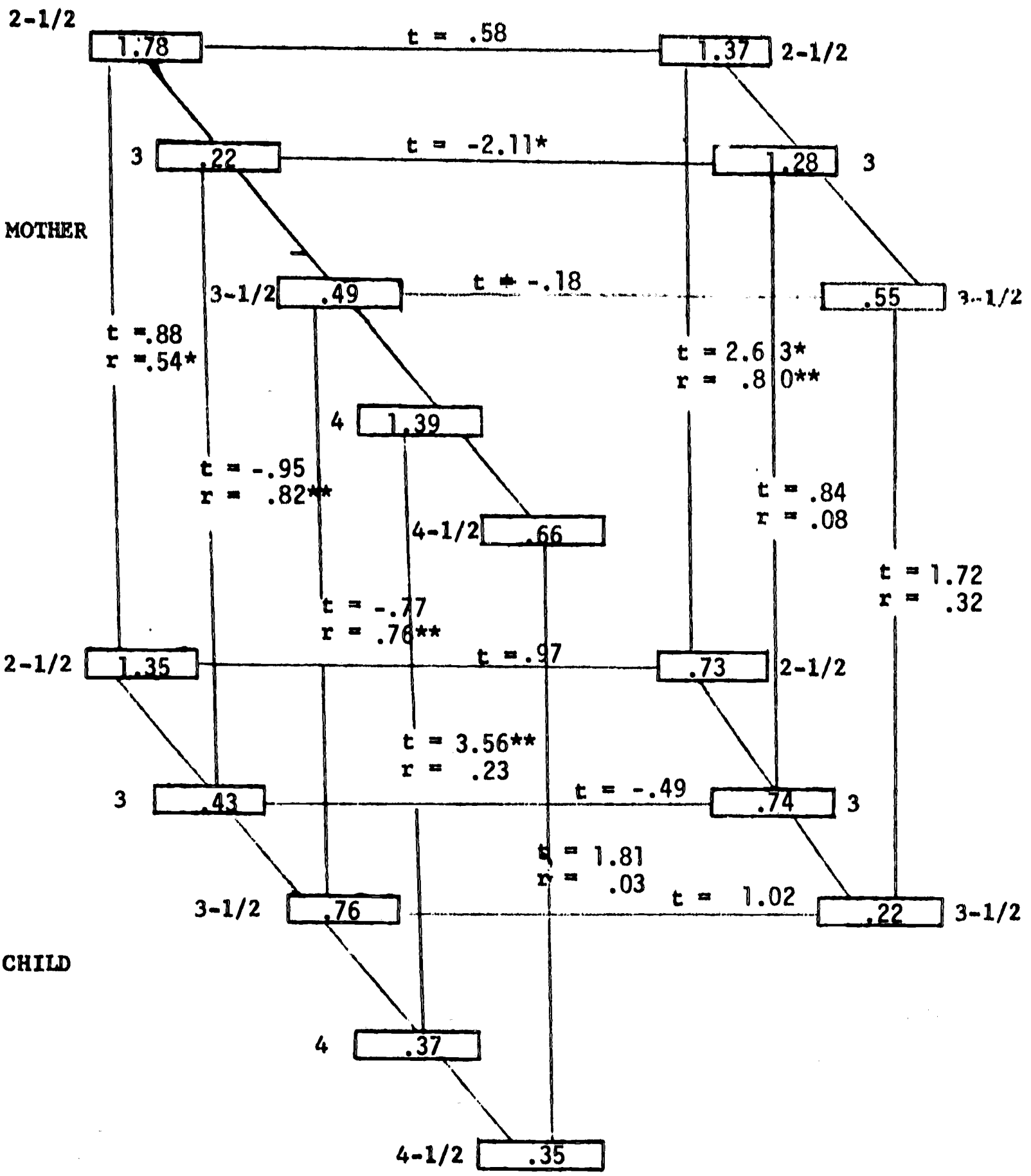


Table 19

Variable: Mode - % Commands

Harlem **Washington Square**

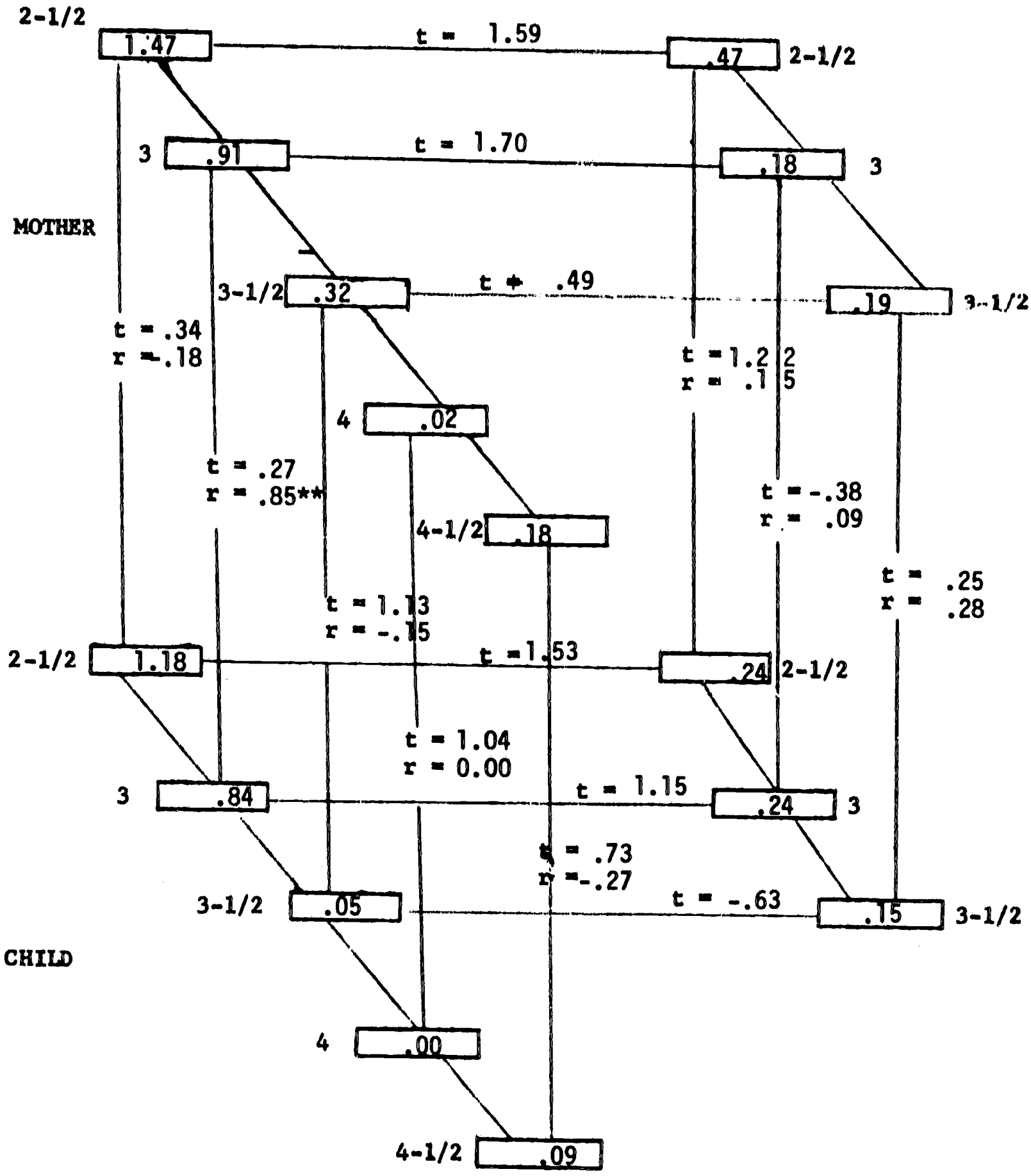


Table 20

Variable: Mode - % Labels

Harlem **Washington Square**

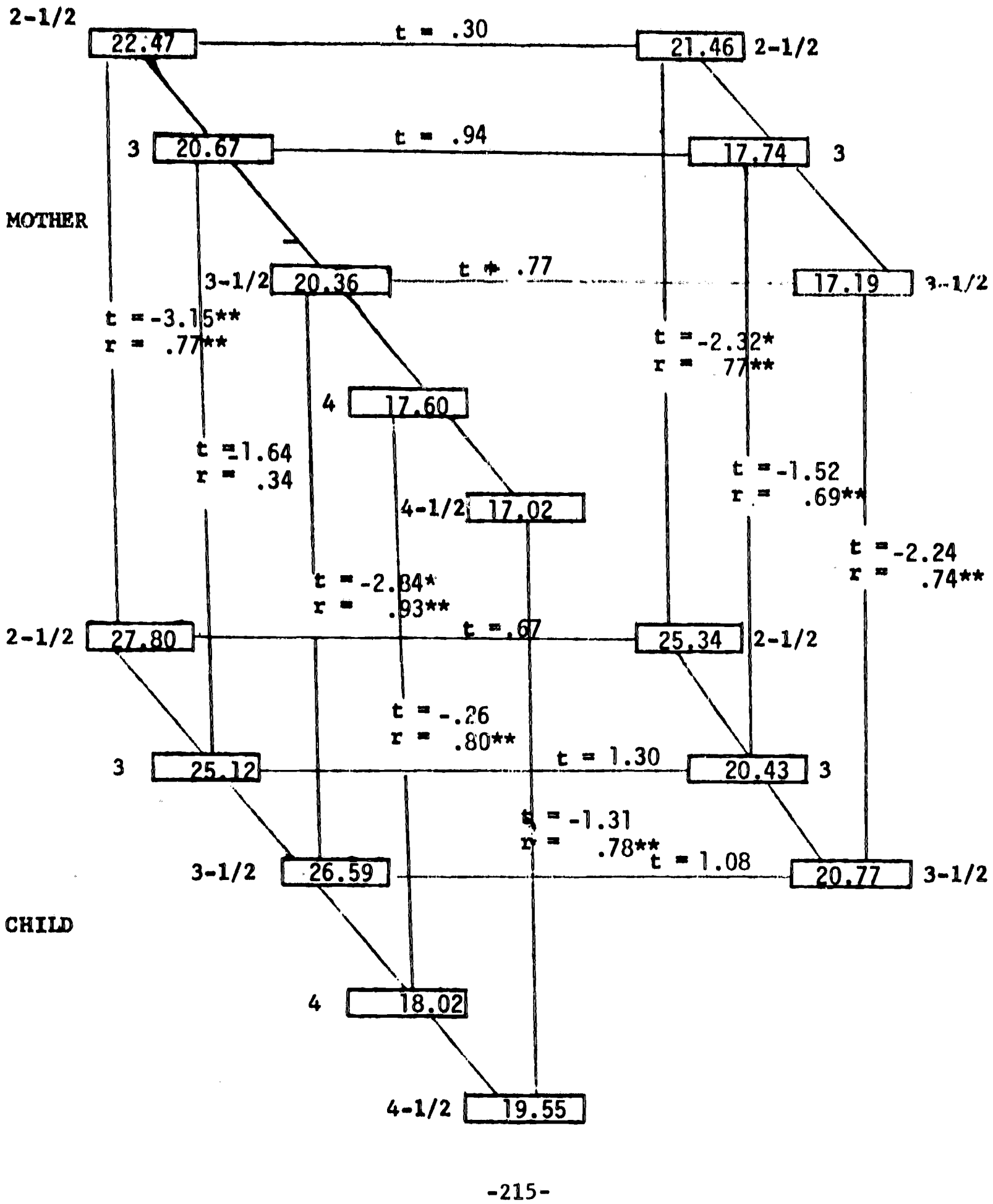


Table 21

Variable: Mode - % Specify

Harlem **Washington Square**

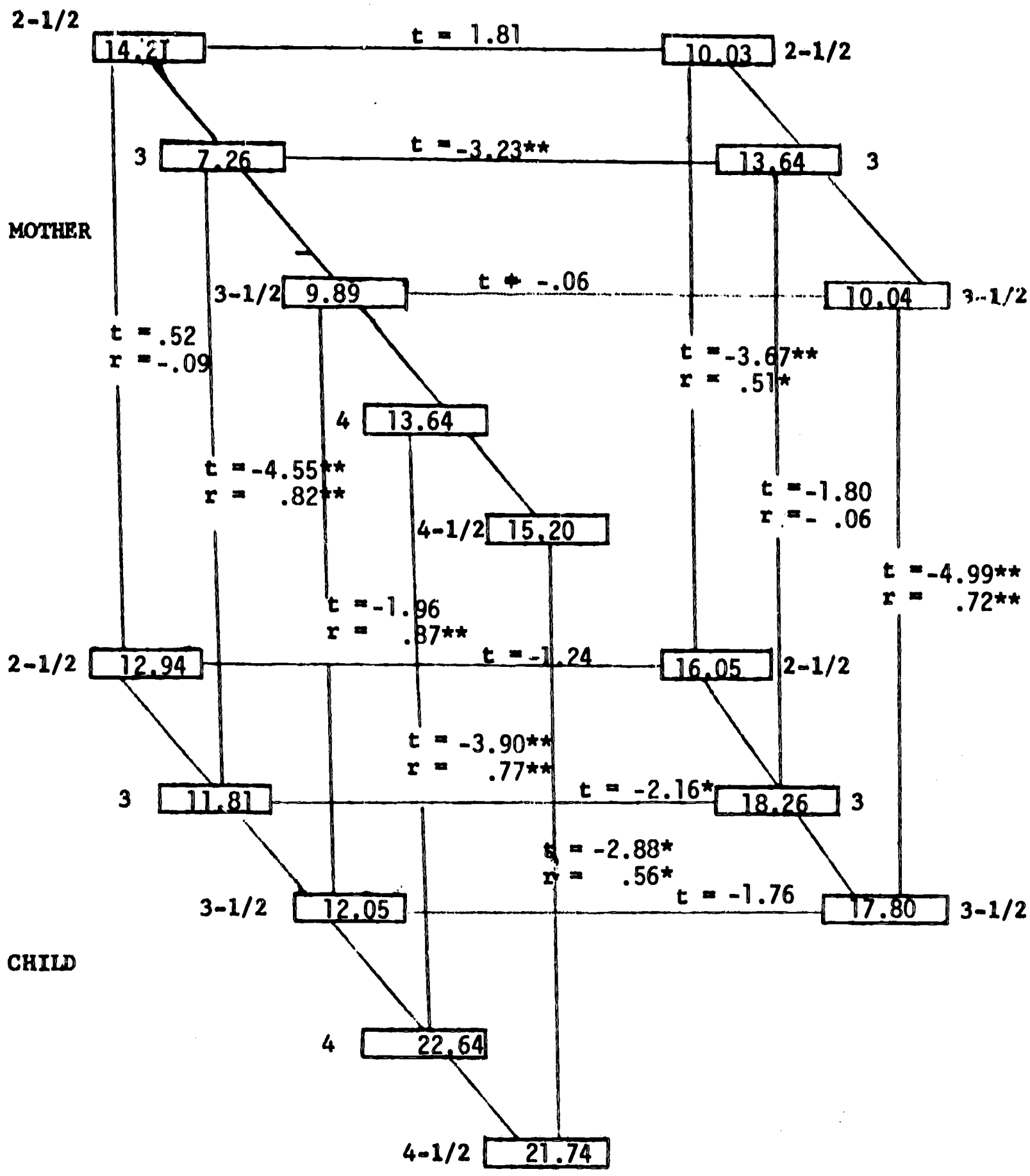


Table 22

Variable: % Utterances Responded To

Harlem

Washington Square

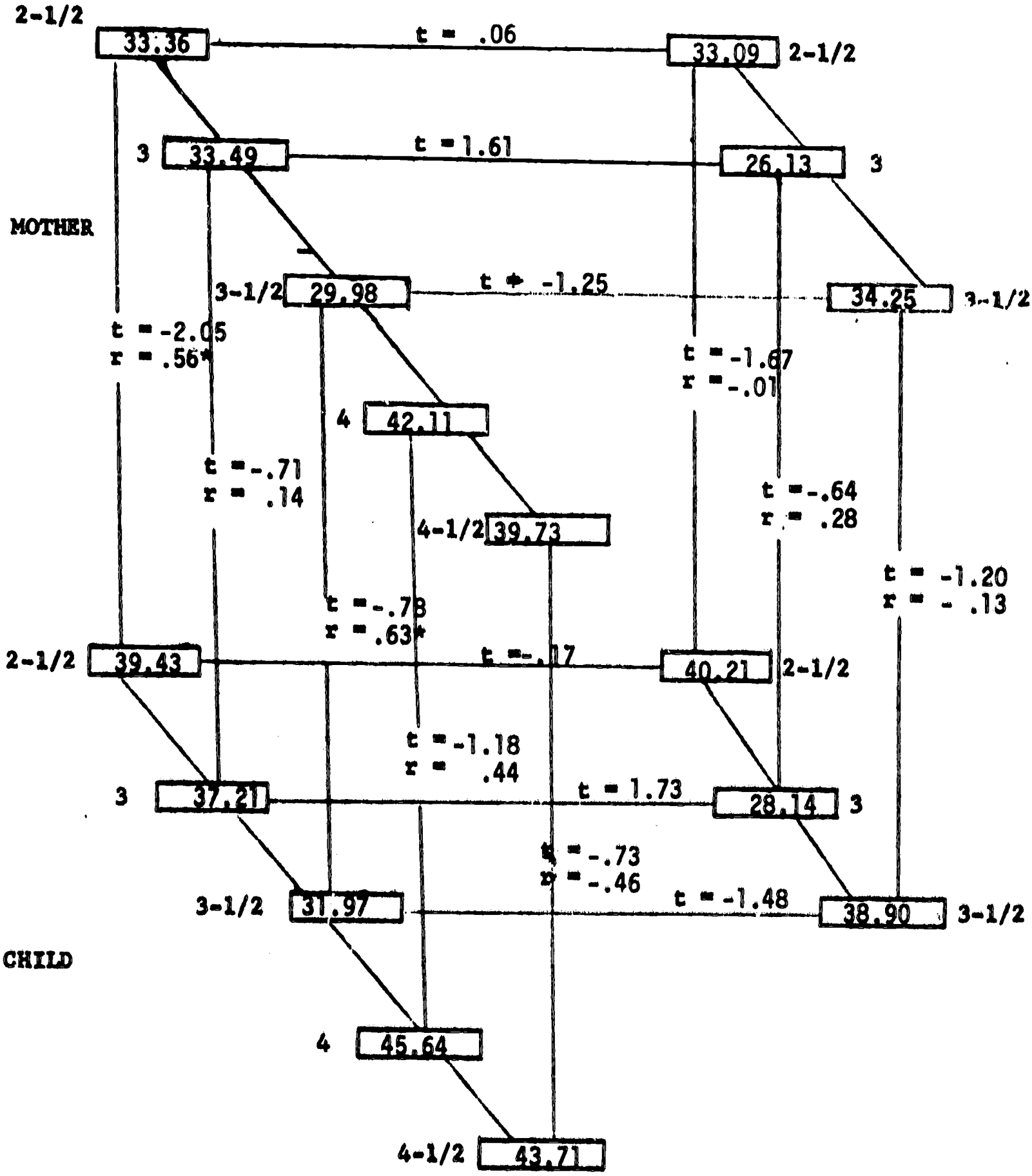


Table 23

Variable: % Utterances Ignored

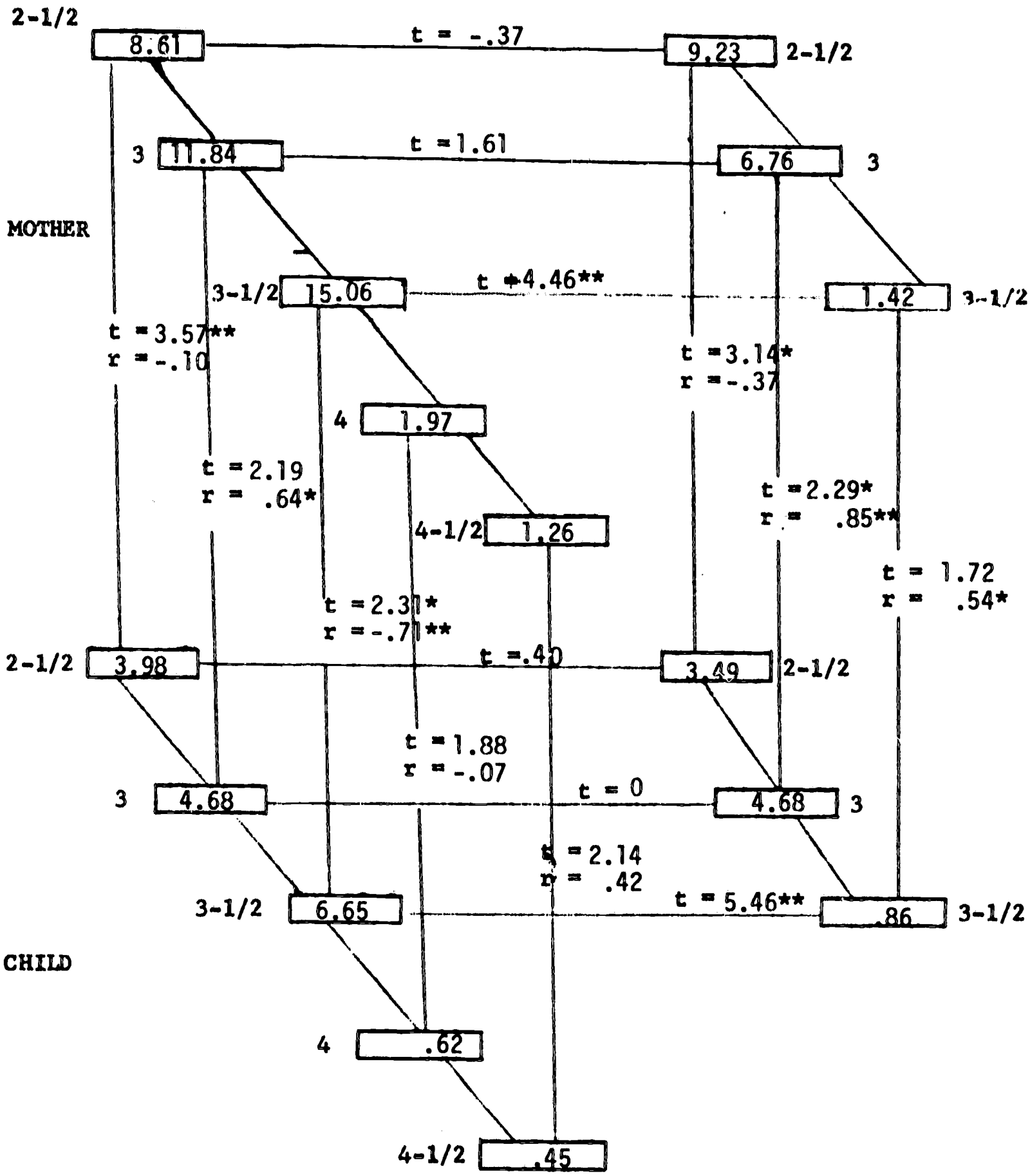


Table 24

Variable: % Statement Containing Clarifying Permanent Information

Harlem **Washington Square**

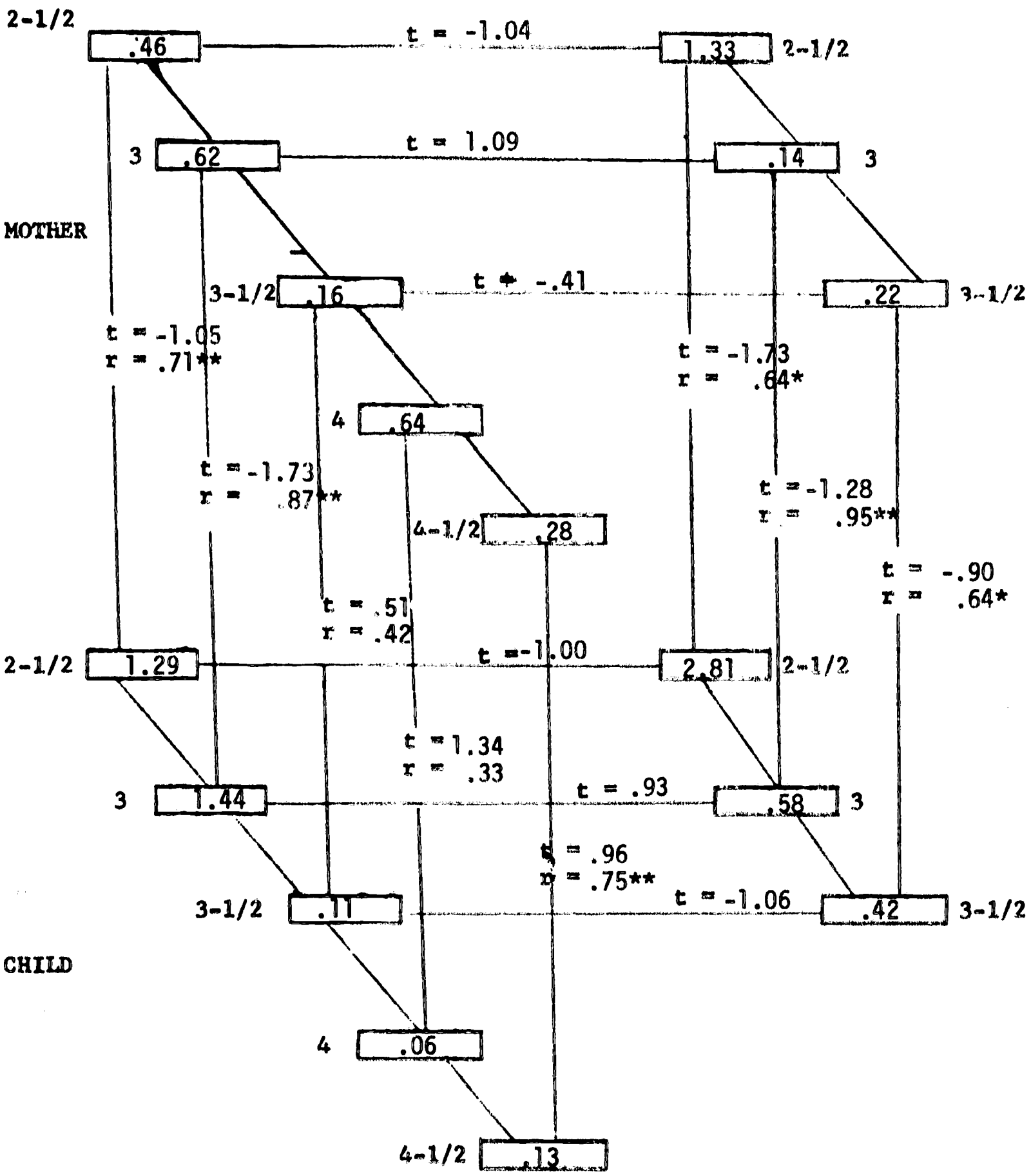


Table 25

Variable: % Statements Containing Clarifying Transient Information

Harlem **Washington Square**

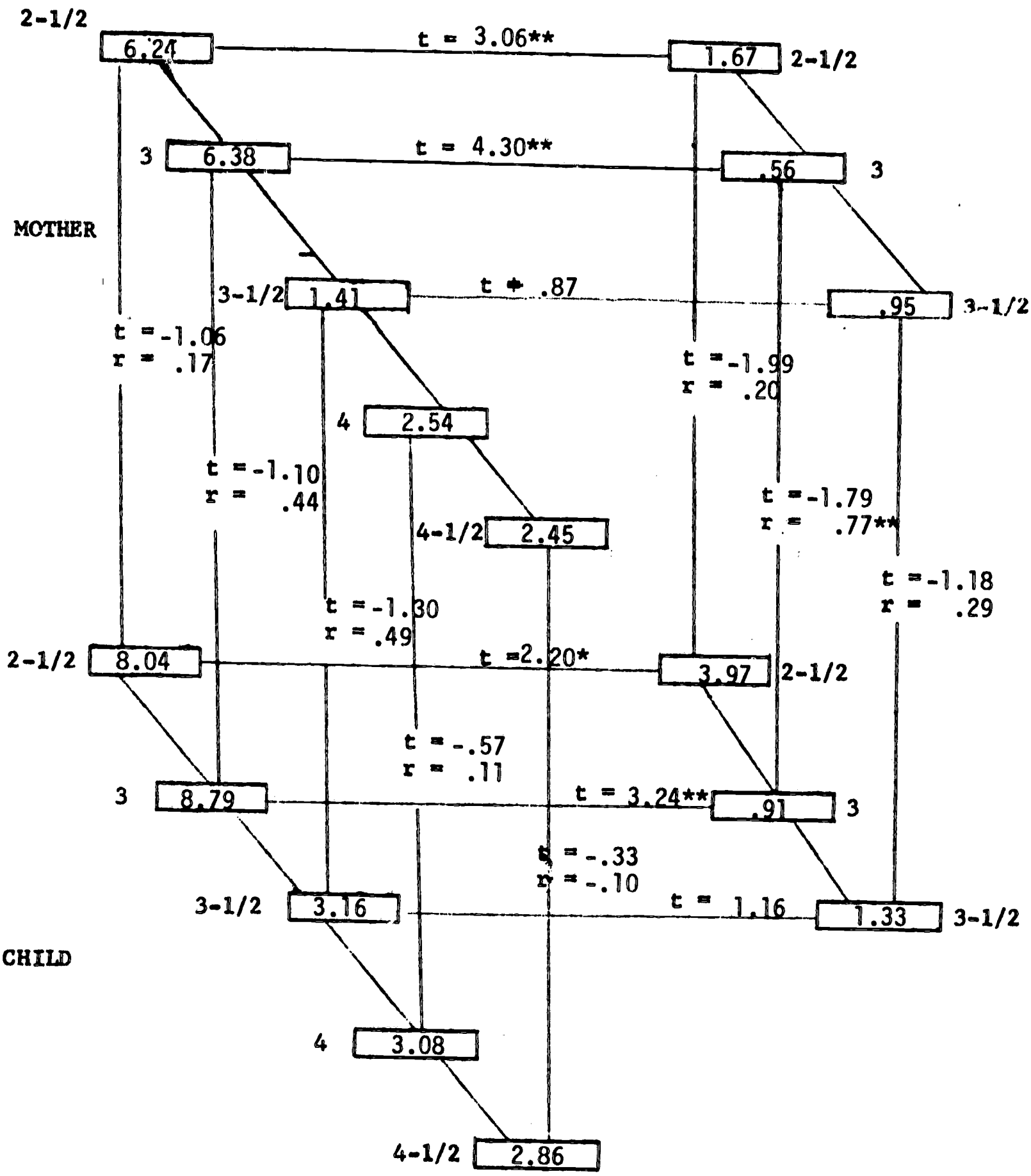


Table 26

Variable: % Questions Directly Responded To

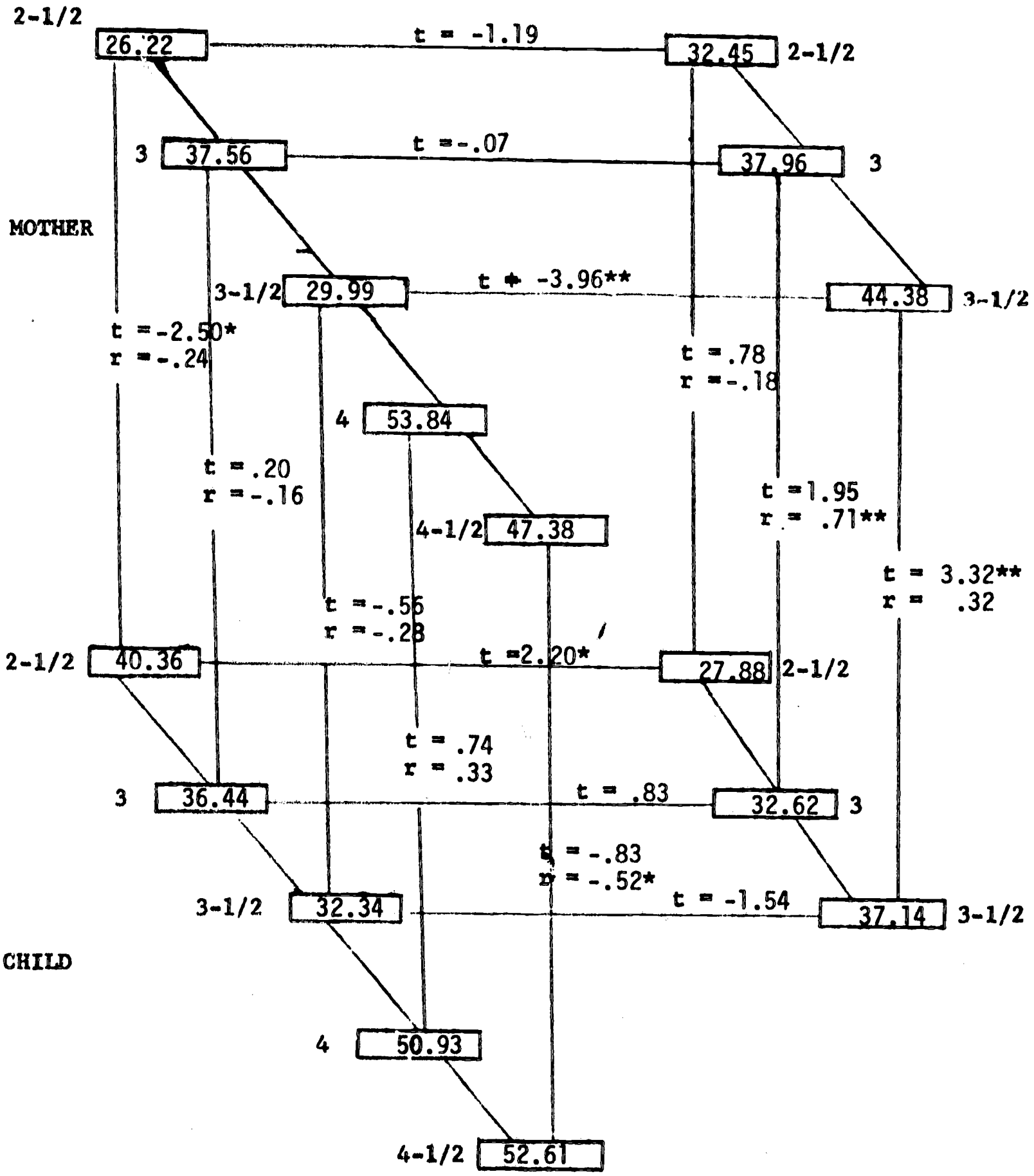


Table 27

Variable: % Questions Peripherally Responded To

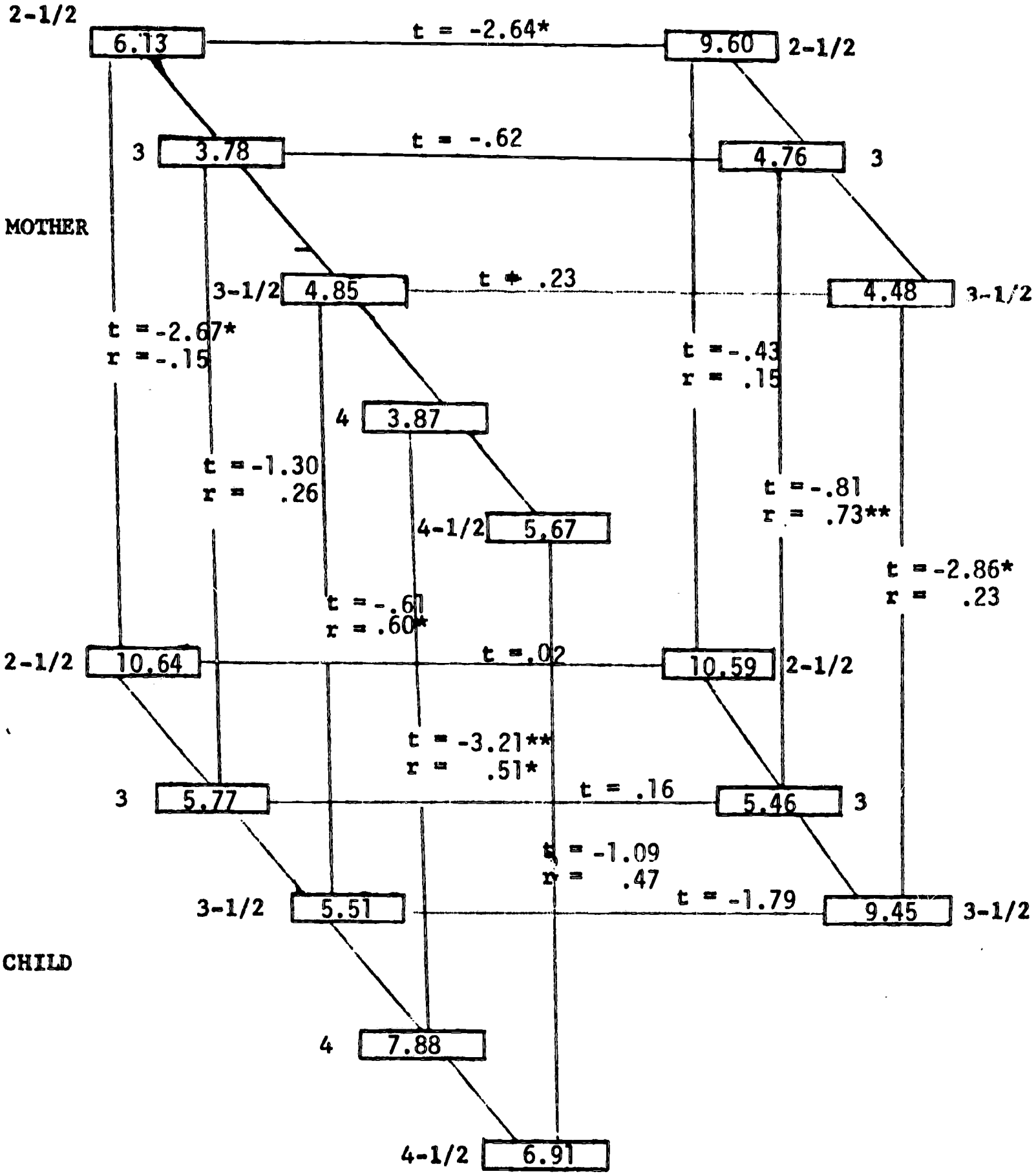


Table 28

Variable: % Questions Responded to With no Information

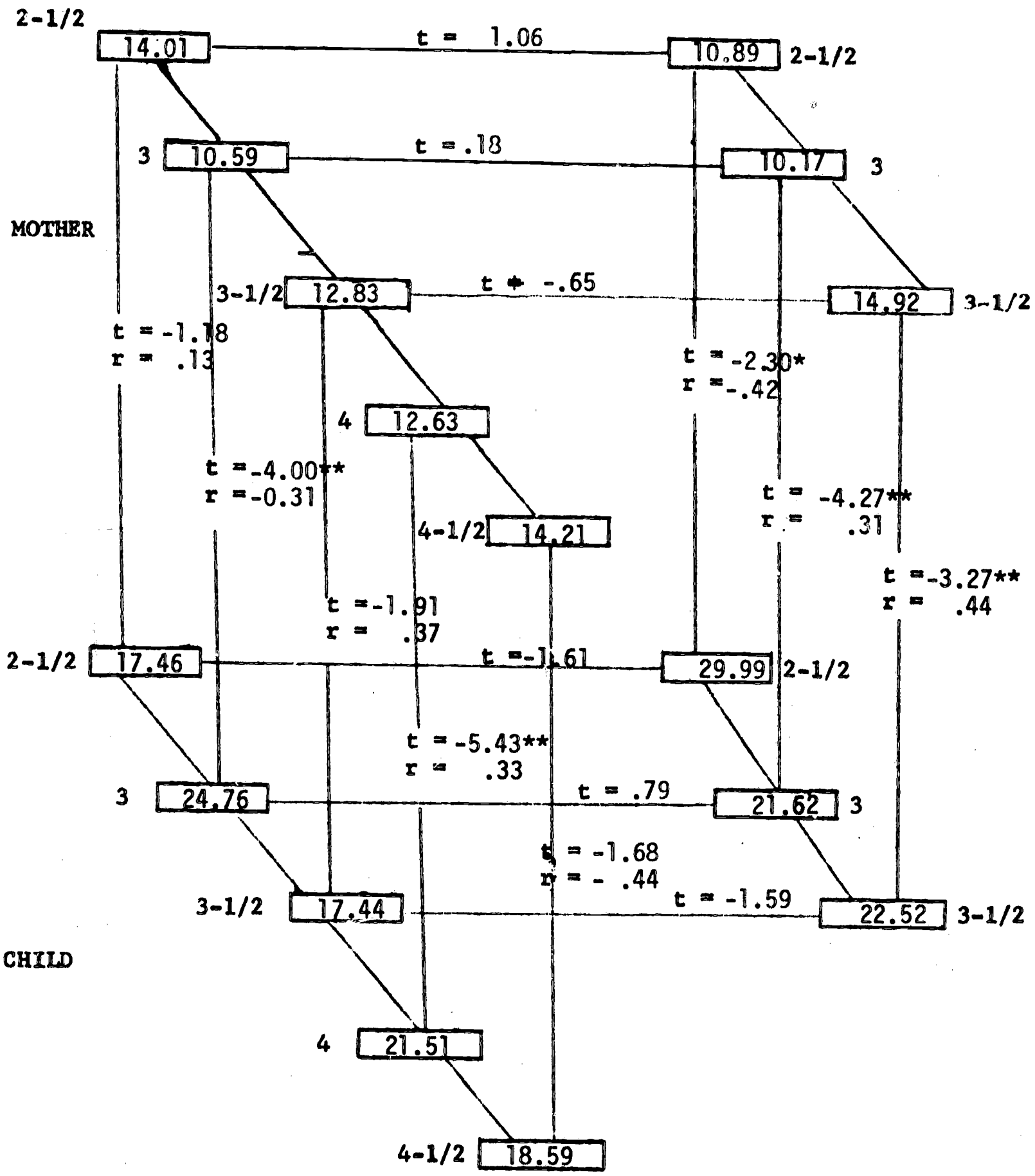


Table 29

Variable: % Statements Directly Responded To

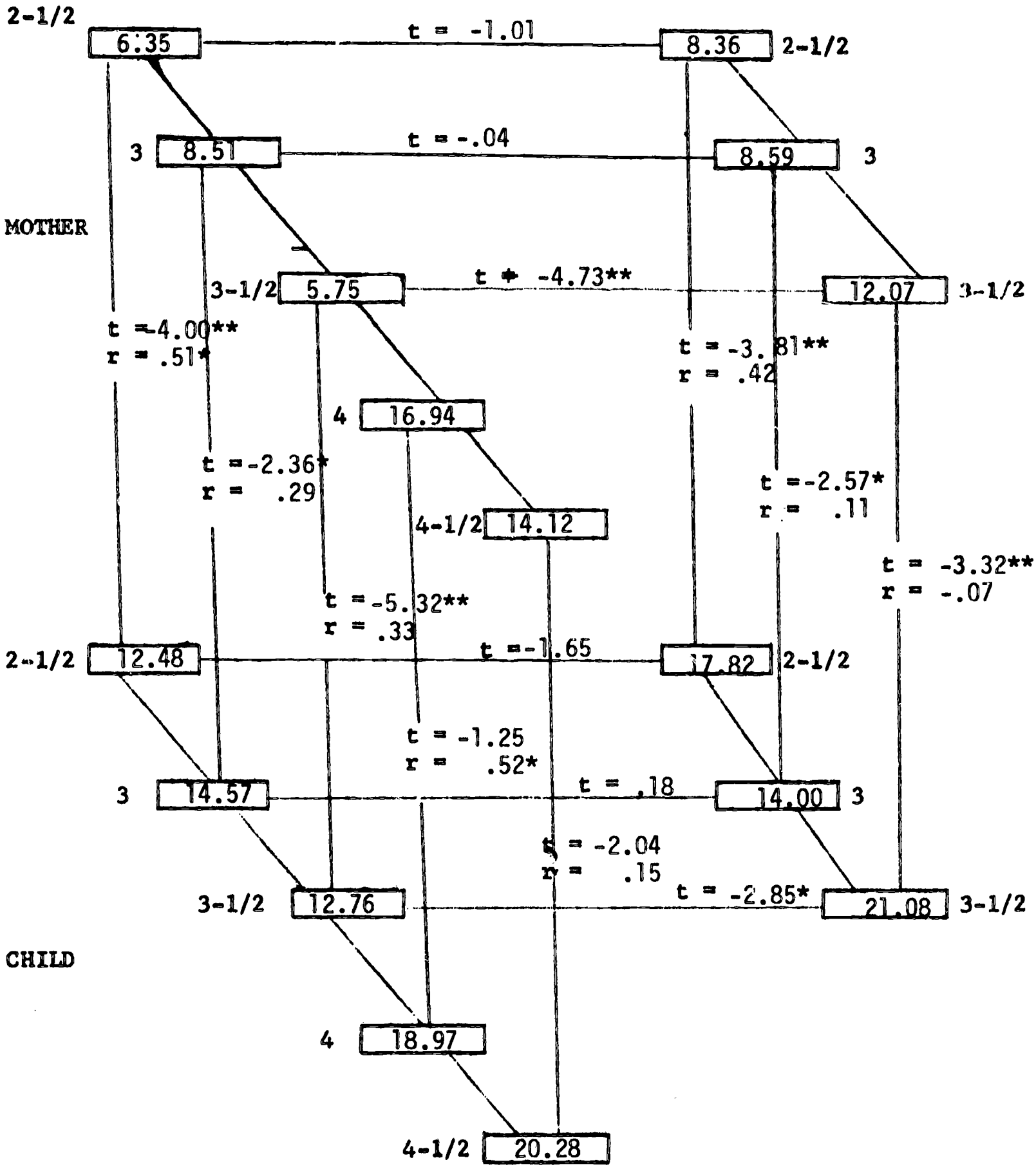


Table 30

Variable: % Statements Peripherally Responded To

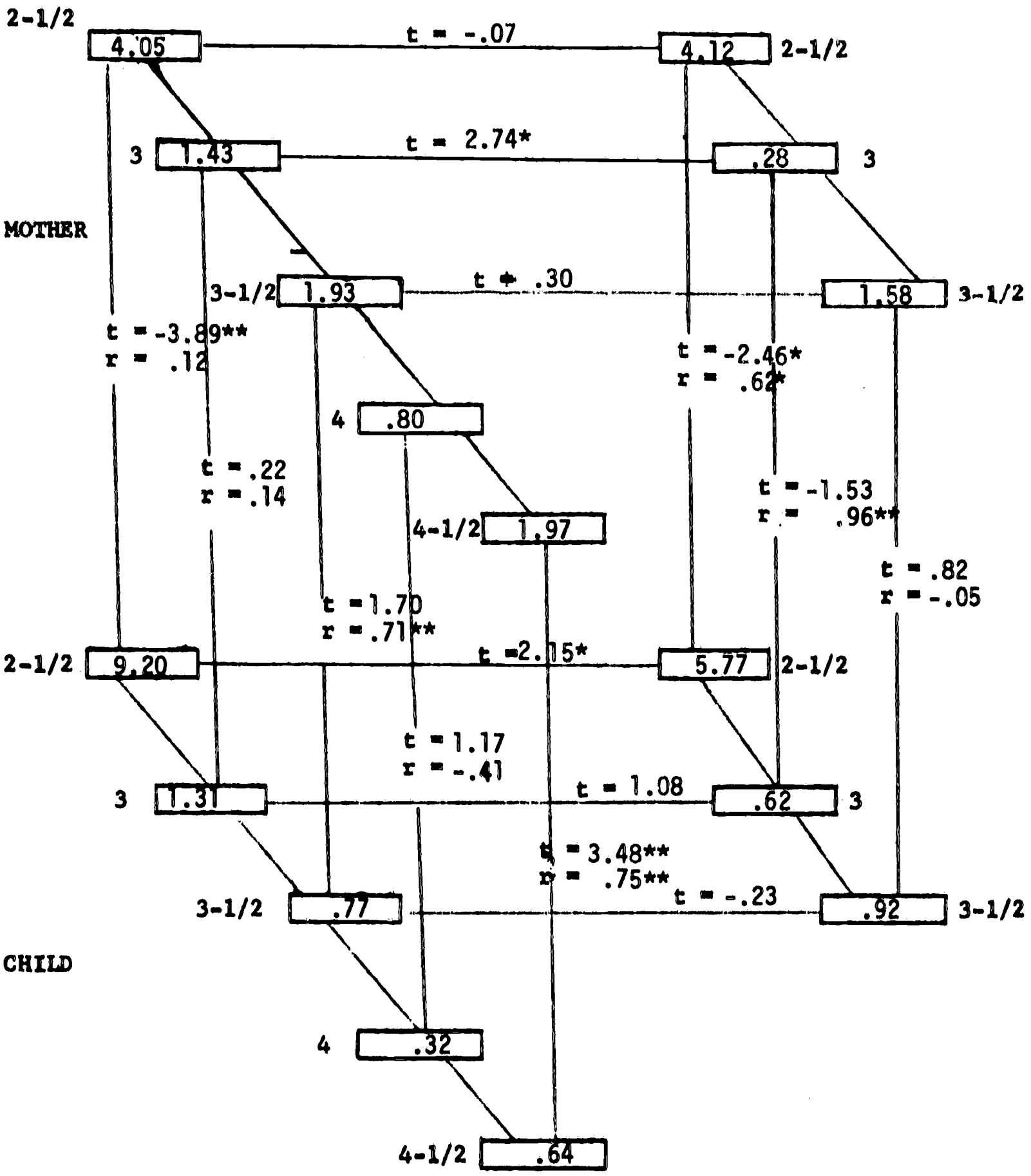


Table 31

Variable: % Statements Responded to with no Information

Harlem **Washington Square**

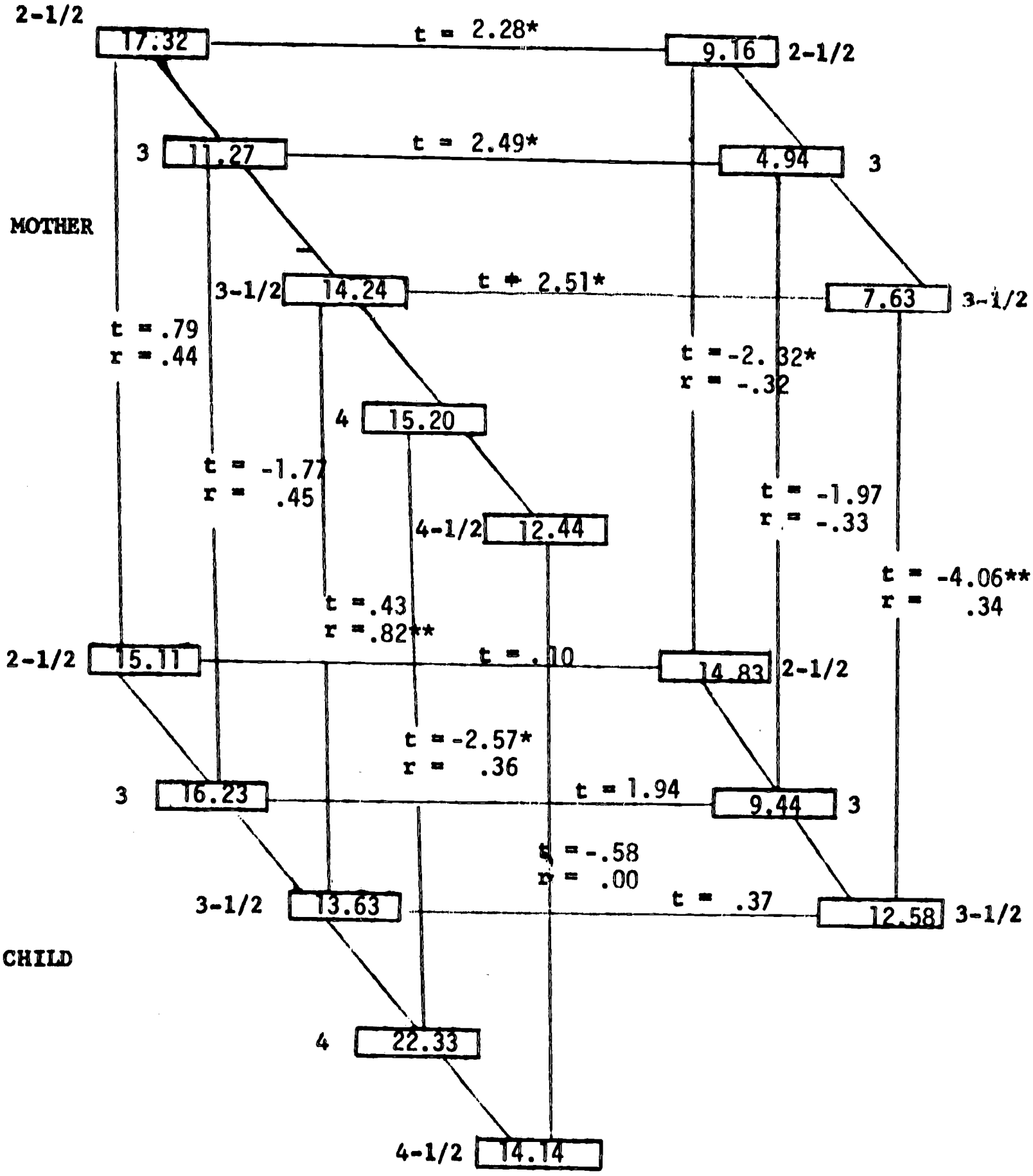


Table 32

Variable: % Questions Requesting Clarifying Transient Information

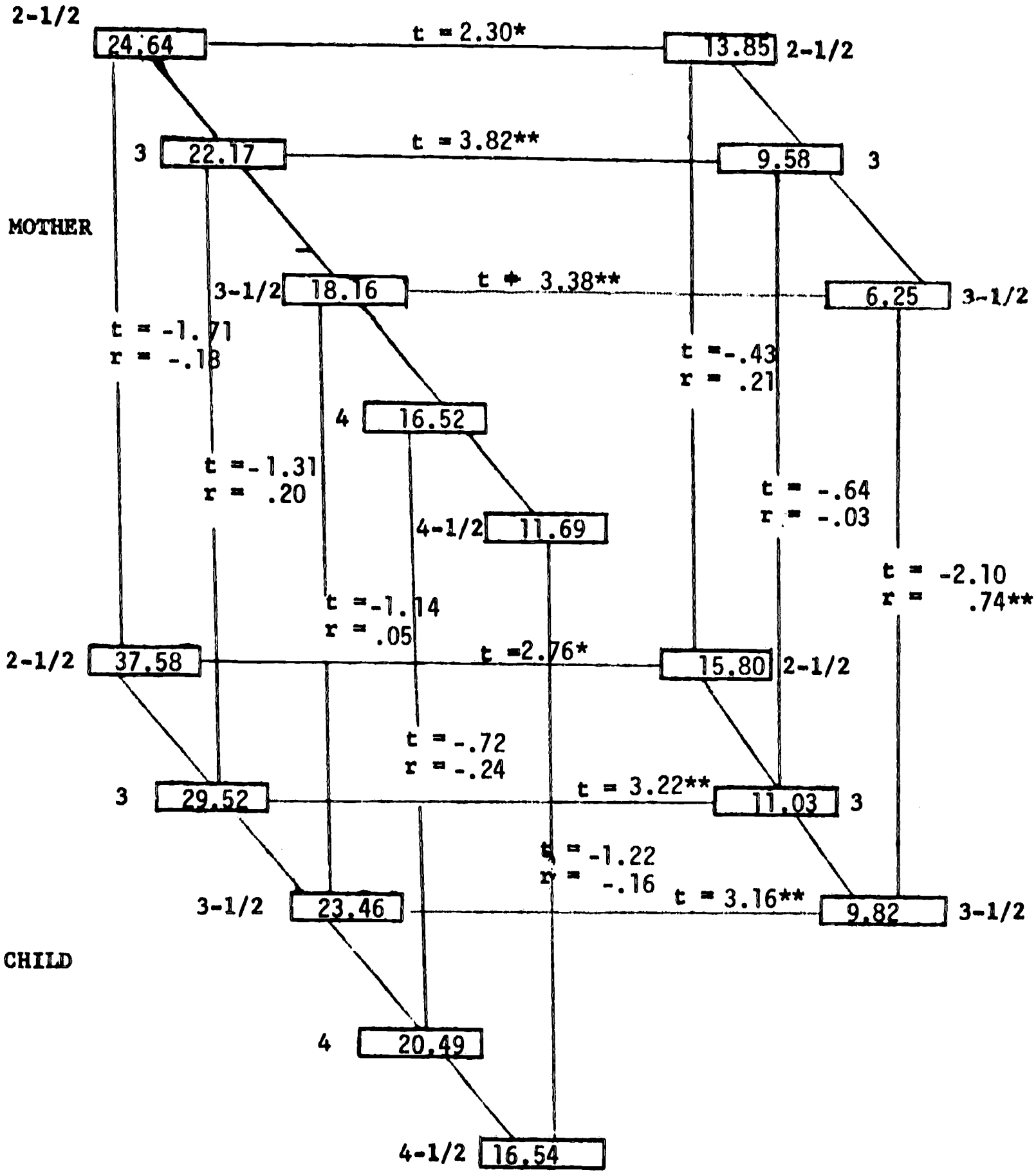


Table 33

Variable: % Questions Requesting Clarifying Permanent Information

Harlem **Washington Square**

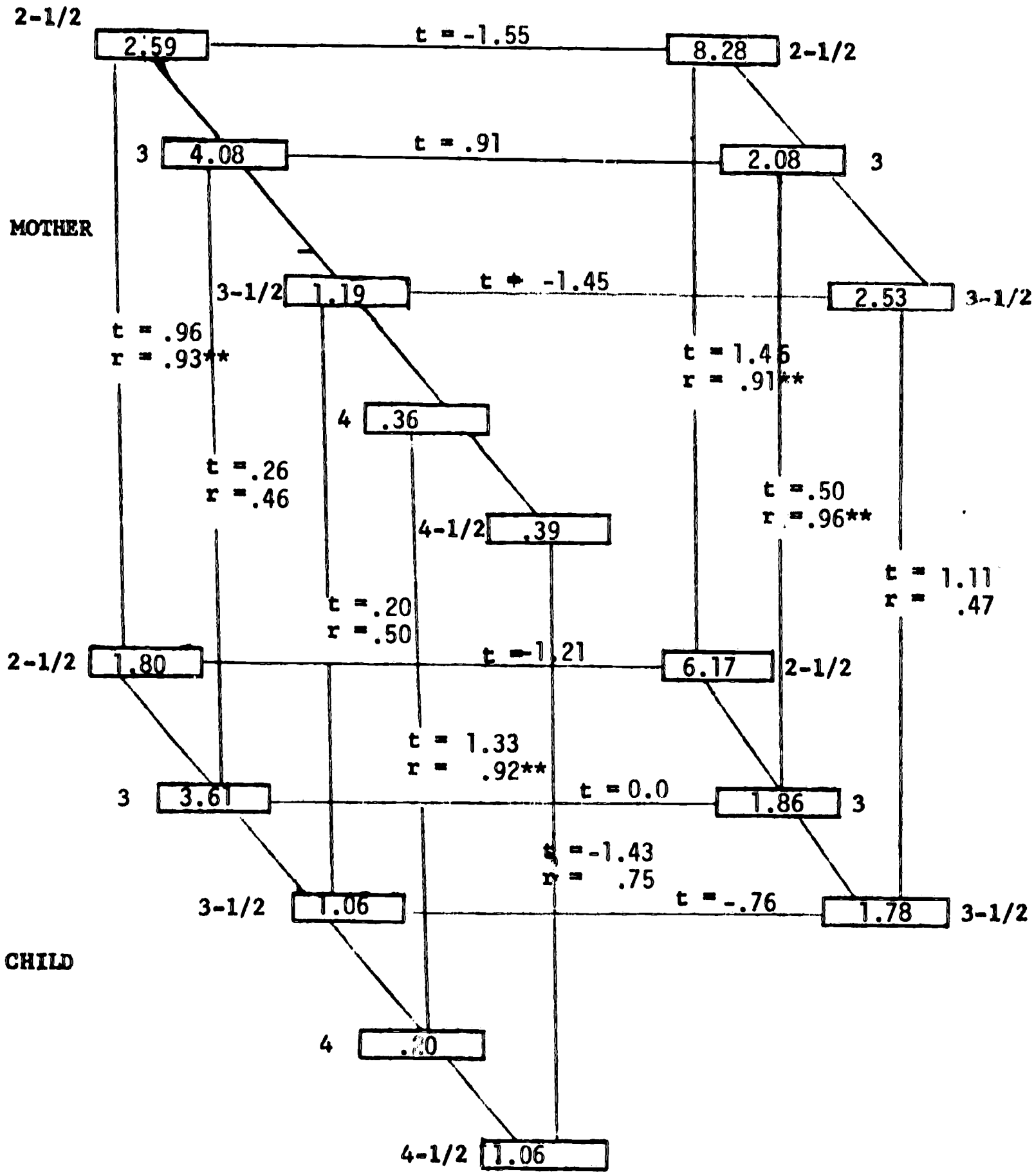


Table 34

Variable: % Utterances Responded to with Direct Information

Harlem **Washington Square**

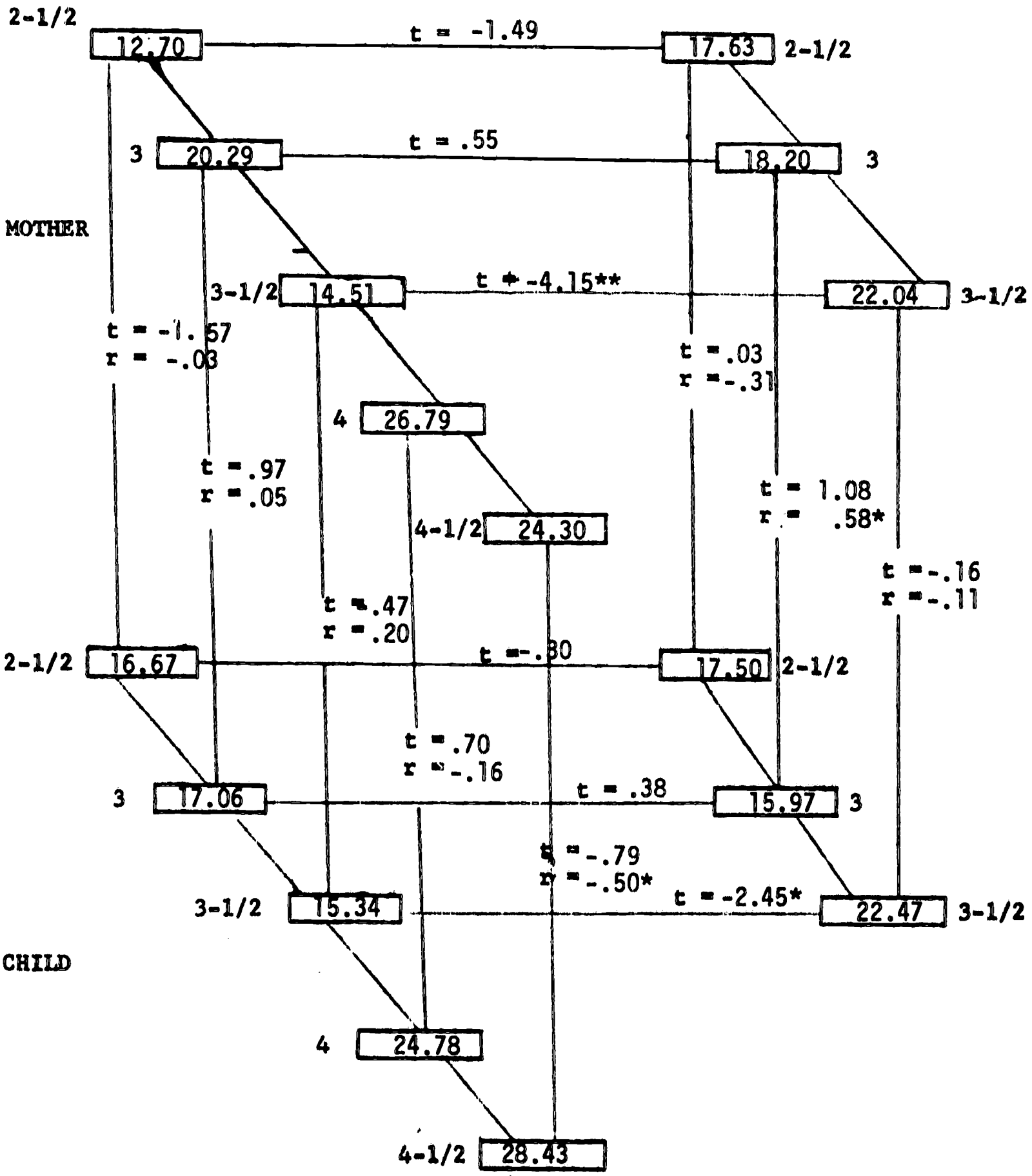


Table 35

Variable: % Utterances Responded to with Peripheral Information

Harlem **Washington Square**

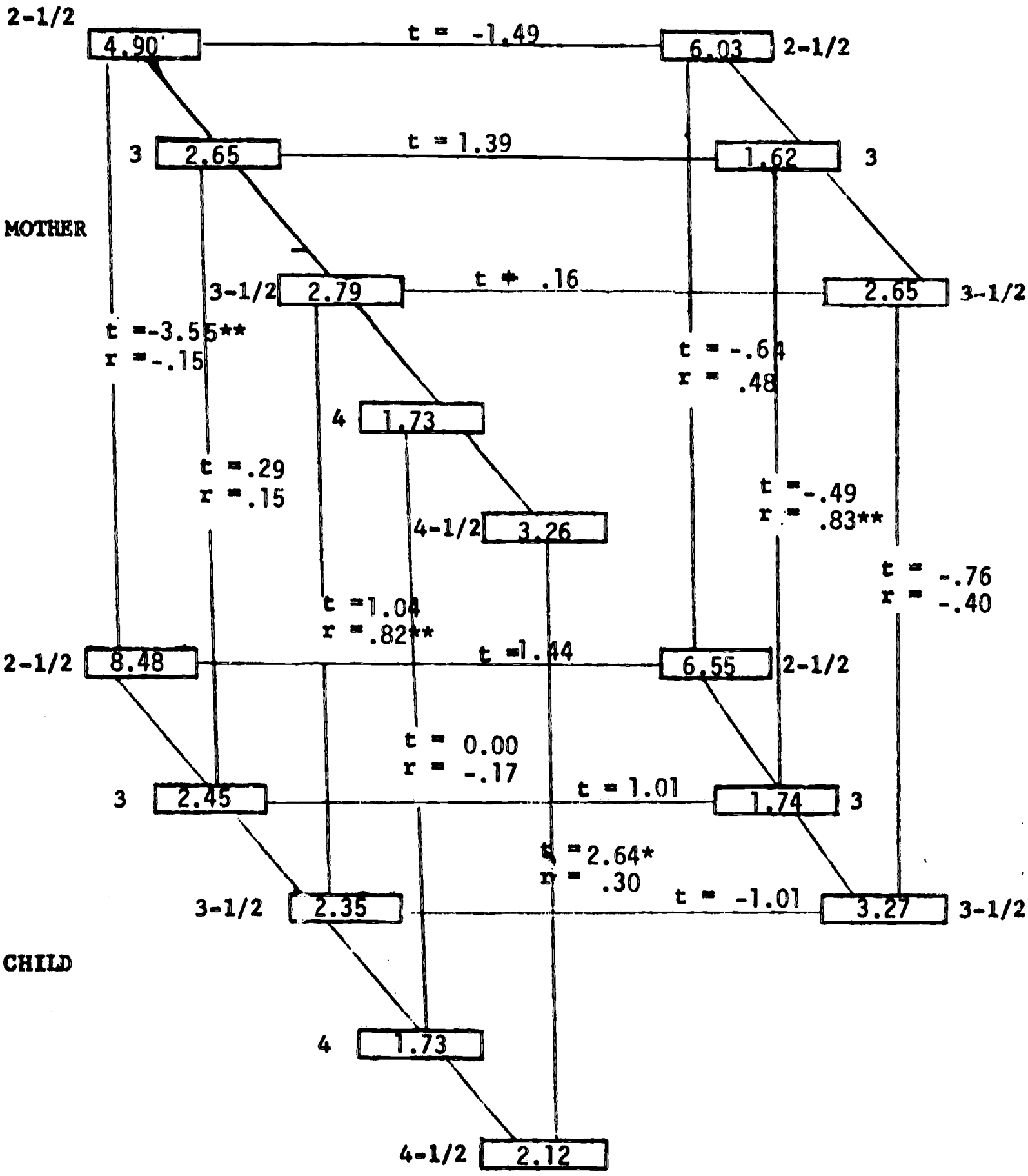


Table 36

Variable: % Coded No Information

Harlem **Washington Square**

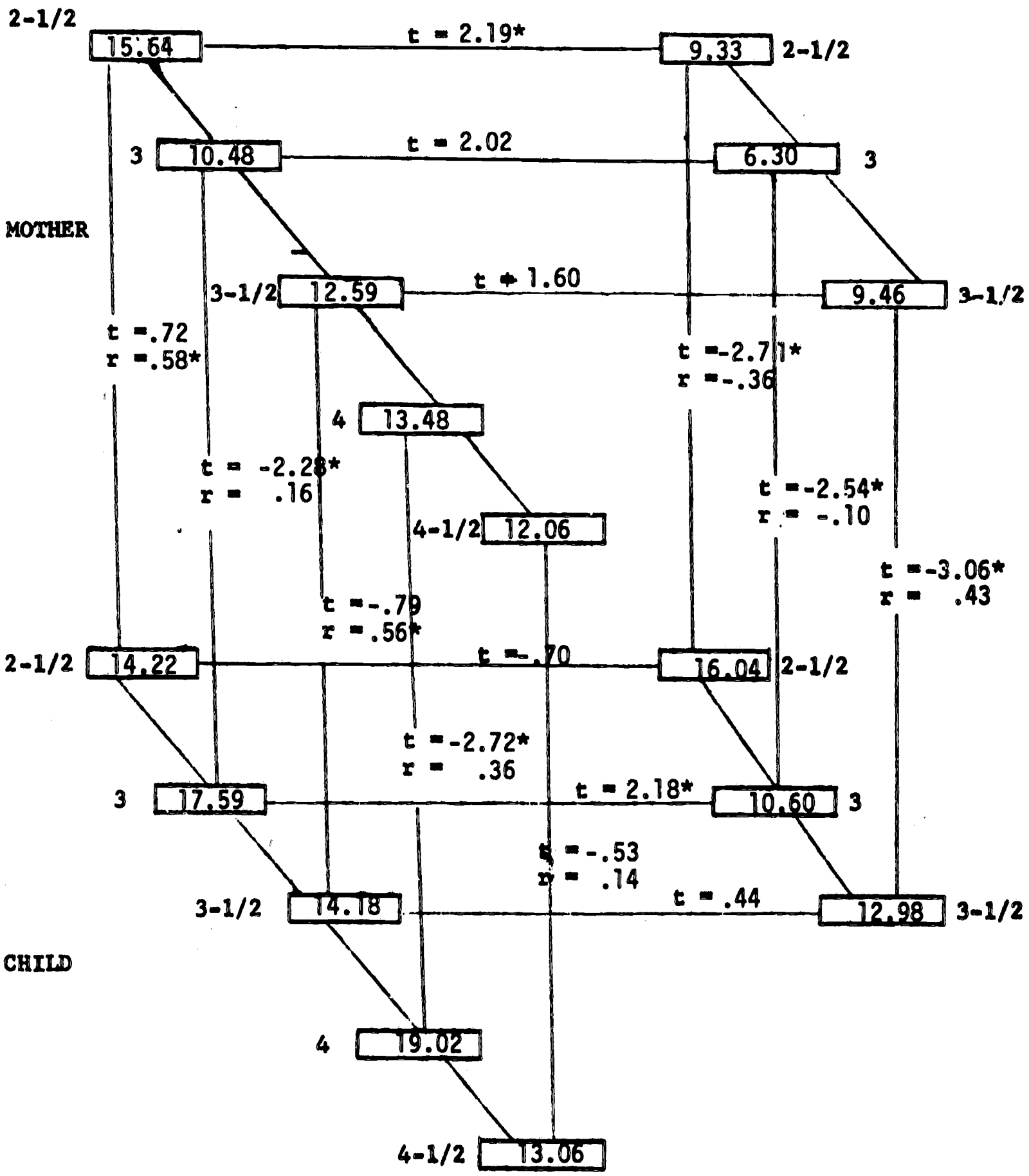
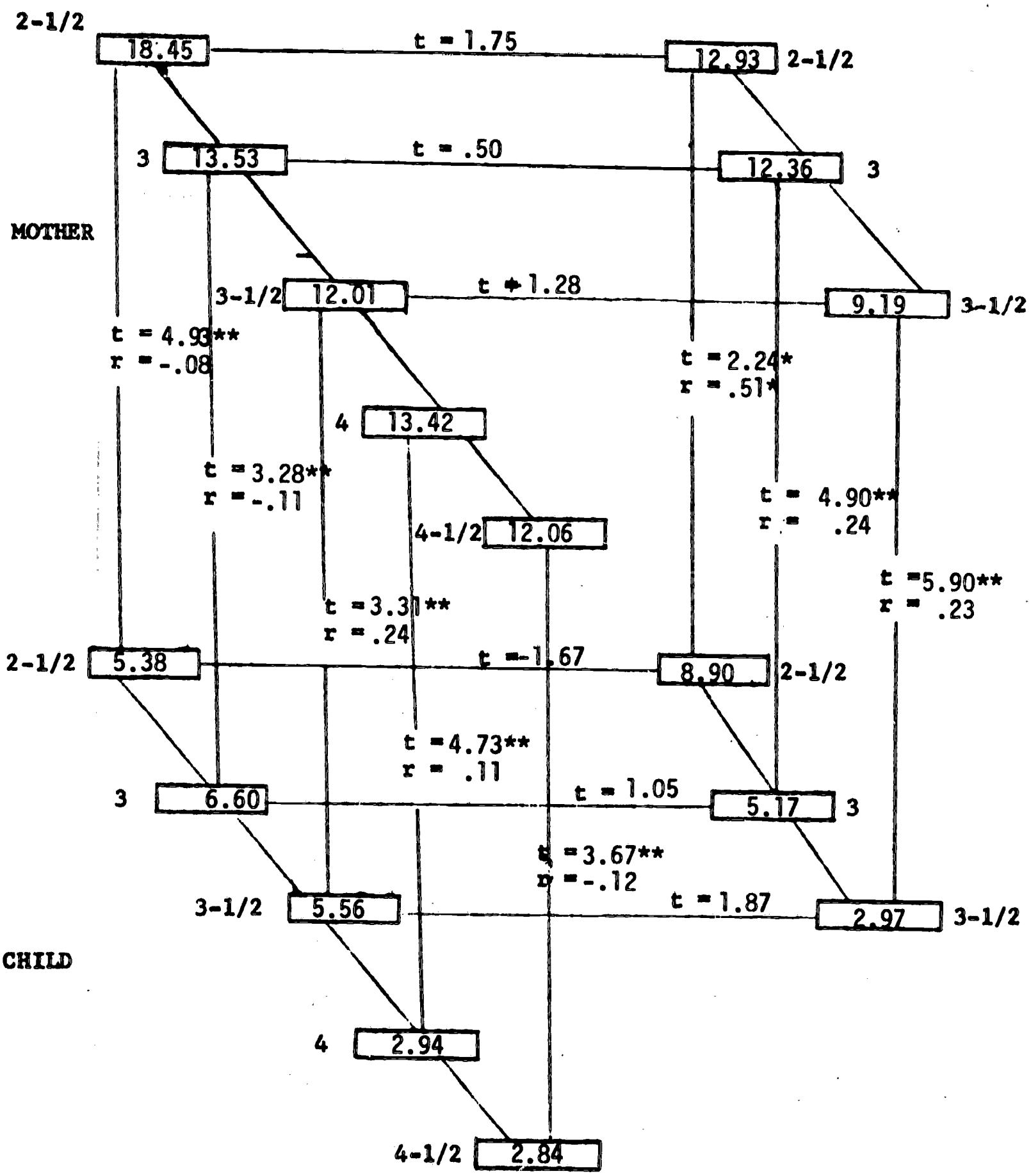


Table 37

Variable: % Behavior Requests

Harlem **Washington Square**



Appendix D

Sample Differences

Table 38

Mean Percentage of Child's utterances falling into VINEX categories

Variable	Session	W. Harlem	Wash. Sq.	Mean diff.	T
Mean No. utterances	1	396.50	306.09	90.41	2.06
	2	344.70	320.73	23.97	.60
	3	296.50	328.20	-31.70	-1.11
Codable utterances percent	1	61.98	80.66	-18.68	-4.45**
	2	76.91	85.57	- 8.66	-2.71*
	3	78.86	79.88	- 1.02	- .28
Form Question Percent	1	34.11	18.53	15.58	3.02**
	2	29.83	23.46	6.37	1.46
	3	29.32	27.94	1.38	.30
Statement Percent	1	65.86	81.42	-15.56	-2.98**
	2	70.01	76.46	- 6.45	-1.47
	3	70.55	71.90	- 1.35	- .30
Information Permanent Percent	1	40.01	42.81	- 2.80	- .59
	2	37.33	28.97	8.36	2.40*
	3	36.32	41.31	- 4.99	- .90
Transient Percent	1	36.72	28.36	8.35	1.78
	2	43.28	47.69	-4.41	- .99
	3	47.52	36.81	10.71	1.92
Fantasy Percent	1	5.49	11.36	-5.86	-1.63
	2	3.51	8.57	-5.06	-2.25*
	3	5.15	9.06	-3.91	-1.30
Behavioral Percent	1	17.32	17.01	.31	.10
	2	15.36	14.51	.85	.37
	3	10.76	12.45	-1.69	-.69
Approval Percent	Too few responses for meaningful analysis				
Disapproval Percent	Too few responses for meaningful analysis				

Variable	Session	W. Harlem	Wash. Sq.	Mean diff.	T
Mode of Information Response					
Explanation					
Percent	1	2.00	3.56	-1.56	-1.11
	2	1.52	2.13	- .61	- .94
	3	1.01	5.36	-4.35	-4.18
Limits					
Percent	1	2.09	2.51	- .42	- .51
	2	.92	1.30	- .38	- .65
	3	4.04	3.96	.08	.06
Clarification					
Percent	1	19.84	10.06	9.78	2.60*
	2	16.77	4.31	12.46	4.23**
	3	9.24	4.70	4.54	2.36*
Describes					
Percent	1	28.79	37.26		-2.00
	2	37.50	45.22		-1.60
	3	42.99	41.13		-0.041
Feeling					
Percent	1	1.81	3.06	-1.26	-2.02
	2	2.96	3.32	- .36	- .40
	3	2.14	2.83	- .69	- .62
Demonstrates					
Percent	1	1.35	.73	.62	.97
	2	.43	.74	- .31	- .49
	3	.76	.22	.54	1.02
Commands					
Percent	1	1.18	.24	.94	1.53
	2	.84	.24	.60	1.15
	3	.05	.15	- .10	- .63
Labels					
Percent	1	27.80	25.34	2.46	.67
	2	25.12	20.43	4.69	1.30
	3	26.59	20.77	5.82	1.08
Specifies					
Direct	1	16.67	17.50	-.83	- .30
	2	17.06	15.97	1.09	.38
	3	15.34	22.47	-7.13	-2.45*
Peripheral					
Percent	1	8.48	6.56	1.93	1.44
	2	2.45	1.74	.70	1.01
	3	2.35	3.27	- .92	-1.01

Variable	Session	W.Harlem	Wash. Sq.	Mean diff	T
no information	1	14.22	16.04	-1.82	-.70
	2	17.59	10.60	6.99	2.18*
	3	14.18	12.98	1.20	.44
child's (?) utterance responded to					
	1	39.43	40.21	- .78	- .17
	2	37.21	28.14	9.06	1.73
	3	31.97	38.90	-6.93	-1.48

* $p < .05$

** $p < .01$

Appendix D

Table 39

Sample Differences

Mean Percentage of Mother's utterances falling into VINEX categories

Variable	Session	W. Harlem	Wash. Sq.	Mean diff	T
Mean No. Utterances					
	1	487.67	384.18	103.48	2.16*
	2	403.60	338.73	64.87	1.10
	3	339.80	350.30	-10.50	-.24
Codeable Utterances					
Percent	1	87.62	90.62	- 3.00	-1.50
	2	91.33	92.40	- 1.07	- .60
	3	91.53	91.61	- .08	- .04
Form Question					
Percent	1	35.42	38.53	- 3.10	-0.66
	2	44.79	35.31	9.48	2.46*
	3	42.63	36.20	6.43	1.36
Statement					
Percent	1	64.48	61.37	3.10	.66
	2	55.10	64.59	- 9.49	-2.47*
	3	57.27	63.67	- 6.40	-1.35
Information Permanent					
Percent	1	36.78	43.90	- 7.12	-1.89
	2	36.01	34.26	1.76	.51
	3	34.42	37.89	- 3.47	- .66
Transient					
Percent	1	27.87	25.76	2.11	0.64
	2	36.86	38.91	-2.05	-.41
	3	39.48	35.71	3.77	.72
Fantasy					
Percent	1	2.96	7.16	-4.20	-1.74
	2	2.82	4.06	-1.24	- .89
	3	4.12	3.45	.67	.47
Behavioral					
Percent	1	31.43	21.07	10.36	2.34*
	2	22.72	21.80	.92	.26
	3	21.12	21.43	- .31	-.12

Variable	Session	W. Harlem	Wash. Sq.	Mean diff.	T
Approval	Too few utterances to be meaningfully coded				
Disapproval	Too few utterances to be meaningfully coded				
Mode of Information Response					
Explanation Percent	1	2.61	3.31	- .70	- .73
	2	2.22	6.24	-4.02	-3.92**
	3	2.83	6.72	-3.89	-2.96**
Limits Percent	1	3.74	3.95	- .20	- .23
	2	2.03	1.58	.45	.60
	3	3.44	3.29	.15	.14
Clarification Percent	1	14.60	10.75	3.86	1.40
	2	16.54	4.90	11.64	4.70**
	3	9.58	4.53	5.05	1.43
Describes Percent	1	35.48	44.50		-2.64**
	2	43.67	52.15		-1.34
	3	50.16	53.47		-0.49
Feeling Percent	1	2.75	2.70	.05	.09
	2	3.60	1.26	2.34	3.69**
	3	1.79	2.25	-.46	- .71
Demonstrates Percent	1	1.78	1.37	.41	.58
	2	.22	1.28	-1.06	-2.11*
	3	.49	.55	-. 06	- .18
Commands Percent	1	1.47	.47	.99	1.59
	2	.91	.18	.73	1.70
	3	.32	.19	.13	.49
Labels Percent	1	22.47	21.46	1.00	.30
	2	20.67	17.74	2.92	.94
	3	20.36	17.19	3.17	.77
Direct Percent	1	12.70	17.63	-4.93	-1.49
	2	20.29	18.20	2.09	.55
	3	14.51	20.04	-7.53	-4.15**

Variable	Session	W. Harlem	Wash. Sq.	Mean diff.	T
Peripheral Percent	1	4.90	6.03	-1.13	-1.49
	2	2.65	1.62	1.03	1.39
	3	2.79	2.65	.14	.16
no information Percent	1	15.64	9.33	6.31	2.19*
	2	10.48	6.30	4.18	2.02
	3	12.59	9.46	3.13	1.60
M's (?) utterances responded to Percent	1	33.36	33.09	.27	.06
	2	33.49	26.13	7.36	1.61
	3	29.98	34.25	-4.27	-1.25
M's utterances ignored Percent	1	8.61	9.23	- .62	- .37
	2	11.84	6.76	5.08	1.65
	3	15.06	1.42	13.64	4.46**

* $p < .05$
** $p < .01$

Appendix D
Table 40
VINEX means and t's by educational level of head of household
W.H. Ages 3,4, 5.

Variable	Age	Mother ME	Mother LE	Mother t	Child ME	Child LE	Child t
Total pair frequency	3	675.22	782.55	-1.45			
	4	638.14	639.46	- .02			
	5	704.43	523.92	2.23*			
(Individual utterance frequency)	3	356	471	-1.89	319	311	.21
	4	326	336	- .15	312	303	.28
	5	371	249	2.97**	334	275	1.15
Coded UH Frequency	3	324	425	-1.78	238	238	0
	4	297	302	- .07	237	243	-.2
	5	334	221	2.89**	268	219	1.17
Form							
Question Frequency	3	122	150	- .79	64	67	-.19
	4	106	93	.58	56	63	-.56
	5	95	66	1.42	59	59	.01
Statement Frequency	3	201	275	-1.81	175	141	.15
	4	192	209	-.38	181	180	.02
	5	239	155	2.92**	209	160	1.43
Percent Coded	3	90.98	89.96	.41	72.34	75.56	-.76
	4	90.19	88.91	.37	75.30	79.17	-.93
	5	90.29	88.57	.43	80.41	77.62	.68

1. (NS - Age 3, 9 ME, 11 LE; Age 4, 9 ME, 13 LE; Age 5, 7 ME, 13 LE)
cut-off point - H.S. graduate

	Age	ME	LE	t	ME	LE	t
Percent Questions	3	37.66	34.44	.54	24.59	28.02	- .74
	4	35.10	30.47	1.23	23.66	27.04	- .69
	5	27.46	27.56	- .02	22.50	27.05	- .83
Percent Statement	3	62.22	65.46	-.55	75.54	71.99	.76
	4	64.89	69.50	-1.22	76.24	72.70	.71
	5	72.44	72.38	.01	77.40	72.82	.83
Percent Behavior Requests	3	13.9	19.38	-1.42	9.73	4.83	3.24**
	4	12.36	17.84	-1.56	2.46	3.55	- .94
	5	14.73	15.47	- .19	3.04	3.52	- .33
Information Content							
Percent Permanent Information	3	37.30	31.35	1.44	30.36	42.09	-2.94**
	4	33.54	33.25	.07	35.49	41.68	-1.03
	5	31.46	35.82	- .88	42.69	40.14	.34
Percent Transient Information	3	35.70	32.71	.61	47.86	37.37	2.26*
	4	32.24	29.22	.68	39.14	33.36	1.01
	5	31.94	30.84	.18	33.57	27.32	.96
Percent Behavioral Information	3	22.43	33.20	-2.22*	17.73	18.86	- .19
	4	26.36	34.95	-2.35*	17.80	20.87	- .81
	5	33.94	27.95	1.61	19.91	20.25	-.09
Percent Fantasy	3	3.23	1.06	2.25*	3.87	1.87	1.87
	4	5.77	1.64	1.68	7.37	3.75	1.13
	5	1.90	4.15	- .91	3.59	12.06	-1.60

Mode	Age	ME	LE	t	ME	LE	t
Percent Explanation	3	3.74	2.56	.95	2.49	1.06	1.56
	4	3.80	3.62	.18	2.01	1.65	.54
	5	3.41	5.08	-1.14	3.89	2.15	1.41
Percent Limits	3	2.57	.68	2.68*	1.11	.26	2.48*
	4	1.67	.67	1.00	1.71	.31	1.41
	5	1.50	1.39	.10	1.14	.81	.38
Percent Clarification	3	12.82	10.83	.58	14.07	12.27	.51
	4	6.96	7.36	- .22	9.70	6.61	1.99
	5	3.76	4.46	- .55	4.11	5.20	- .59
Percent Description	3	44.24	45.86	- .38	42.23	34.83	1.58
	4	49.36	49.86	- .11	44.59	40.38	.78
	5	54.36	49.42	.97	43.49	43.05	.06
Percent Feeling	3	3.94	2.46	2.14*	3.14	3.17	- .02
	4	1.21	1.47	- .42	1.56	1.90	- .40
	5	2.11	2.06	.08	1.87	3.23	-1.17
Percent Demonstration	3	.72	1.39	-1.30	.51	.08	1.46
	4	1.67	1.41	.56	.33	.38	- .21
	5	1.04	1.59	- .63	.81	.24	1.88
Percent Command	3	.11	1.24	-2.33*	.24	.68	- .79
	4	.03	.77	-1.16	0	0	0
	5	.03	.45	-1.19	0	.22	- .91
Percent Labels	3	20.88	20.39	.13	18.29	30.71	-2.87**
	4	19.09	17.90	.29	19.49	20.69	- .26
	5	15.73	19.43	-1.05	18.51	23.25	- .95

		Mother			Child		
	Age	ME	LE	t	ME	LE	t
Percent Specifies	3	8.63	12.20	-1.26	14.56	14.62	- .02
	4	14.20	14.61	- .14	17.27	25.02	-1.88
	5	16.16	14.15	.61	24.17	18.55	1.45
Response							
Percent Responded to	3	35.26	30.93	.98	33.80	44.82	-1.77
	4	42.74	38.20	1.14	41.61	39.41	.40
	5	37.19	36.78	.11	38.57	32.96	2.03*
Percent Ignored	3	9.54	7.60	.50	4.78	2.78	.89
	4	2.04	2.65	- .52	.61	1.45	1.04
	5	1.13	1.95	- .94	1.11	1.42	- .77
Utterances responded to with direct information	3	21.54	20.34	.36	16.66	25.01	-2.52*
	4	25.79	23.20	.75	21.01	22.67	- .55
	5	23.57	21.70	.60	21.89	20.28	.57
Percent utterances responded to with peripheral information	3	2.54	1.56	1.32	2.90	1.80	1.81
	4	3.30	2.17	1.31	2.81	2.13	.88
	5	3.21	2.80	.43	2.84	2.53	.41
Percent of utterances responded to with no information	3	11.06	8.96	.99	14.04	17.84	- .94
	4	13.57	12.75	.44	17.67	14.54	.93
	5	10.27	12.04	- .76	13.59	10.05	1.06
Percent Questions with Permanent clarifying information	3	3.89	.69	1.70	4.30	1.16	1.68
	4	.97	.44	.70	.76	.17	1.15
	5	.80	.41	.78	.36	.60	- .38

		Mother			Child		
	Age	ME	LE	t	ME	LE	t
Question with transient clarifying information	3	19.41	22.45	- .65	31.02	22.36	1.36
	4	13.47	18.02	-1.03	30.11	15.91	2.51*
	5	9.63	12.23	- .57	15.69	14.54	.21
Statement with permanent clarifying information	3	.67	.40	.46	1.40	.33	1.26
	4	.04	.55	- .91	.04	.05	- .14
	5	.10	0	2.13	0	.08	- .72
Statement with transient clarifying information	3	5.49	2.33	1.79	4.83	7.58	- .93
	4	2.99	1.62	1.39	3.00	3.16	- .11
	5	1.36	1.84	- .51	1.00	1.52	- .95
Question directly responded to	3	40.99	41.10	- .02	40.81	39.76	.21
	4	53.51	48.04	.94	47.23	45.27	.38
	5	54.10	43.64	1.54	39.34	44.72	1.15
Question peripherally responded to	3	4.36	2.78	1.01	6.99	6.13	.44
	4	5.27	4.15	.62	7.07	9.91	- .83
	5	6.01	5.35	.31	11.50	9.01	.72
Question responded to with no information	3	12.77	10.48	.82	20.68	26.21	-1.29
	4	13.81	15.41	- .56	22.59	18.20	1.32
	5	13.91	15.13	- .37	18.77	17.98	.15

		Mother			Child		
	Age	ME	LE	t	ME	LE	t
Statement directly responded to	3	12.04	12.77	- .25	13.88	23.57	-2.68*
	4	14.90	15.31	- .14	18.53	17.63	.36
	5	14.56	15.37	- .34	20.29	16.56	1.17
Statement peripheral responded to	3	1.08	1.10	- .05	2.50	.69	2.00
	4	2.17	1.58	.55	1.09	.41	1.92
	5	2.23	1.92	.28	1.14	.46	1.38
Statement responded to with no information	3	11.43	9.16	.84	15.09	17.31	- .52
	4	14.26	13.10	.45	20.51	16.15	.96
	5	9.90	12.59	- .89	13.81	9.49	1.13

Appendix D
Table 41

Age Comparisons - Children

West Harlem (WH), Harlem Cross-Sectional (CS) and Washington Square (WS)
Child Means and Mean Differences[Q] for Age Comparisons on VINEX Variables

		AGE						
Variable	Group	2½	3	3½	4	4½	5	Additional Significant Differences
Individual Utterance Frequency	WH	397.56	353.11 *	287.44	300.78	305.33		2½-3½*
	CS		314.90		299.65		295.40	
	WS	307.60	319.30	328.20				
Coded UH Frequency	WH	257	277.44 *	225.78	236.11	250.22		
	CS		238.40		233.90		235.95	
	WS	247.40	276.20	263.3				
Question Frequency	WH	87.44	81.89	67.56	58.22	67.89		2½-3½**
	CS		65.45		55.60		58.75	
	WS	45.80	63.80	75.20				
Statement Frequency	WH	169.56	195.56 *	158.11	177.89	182.33		
	CS		173.25		178.15		177.15	
	WS	201.70	212.3	187.9				
Percent Coded	WH	63.39 **	76.67	77.96	77.67	81.51		2½-3½** 2½-4½**
	CS		74.12		76.84		78.60	
	WS	80.1	85.12 *	79.88				
Percent Questions	WH	32.53	27.51	28.33	24.56	26.73		
	CS		26.48		24.30		25.46	
	WS	17.64	22.77	27.94				2½-3½**

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Statement	WH	67.46		72.32		71.53		75.34	73.17		
	CS			73.59				75.50		74.42	
	WS	81.42		76.46		71.90					2½-3½**
Percent Behavior Requests	WH	6.14		6.82		5.17	*	2.20	2.50		3½-4½** 2½-4½*
	CS			7.04		**		3.02		3.35	3-5**
	WS	9.24		5.36	**	2.97					2½-3½**
Percent Permanent Information	WH	40.30		38.09		37.19		38.44	40.06		
	CS			36.81				38.89		41.03	
	WS	41.36	*	29.49	**	41.31					
Percent Transient Information	WH	33.76		41.52		47.10		37.71	37.2		2½-3½**
	CS			42.09				36.10		29.51	3-5**
	WS	29.05	**	46.54	*	36.81					
Percent Behavioral Information	WH	18.06		16.34	*	10.38	*	17.50	17.72		2½-3½* 3½-4½**
	CS			18.08				19.70		20.14	
	WS	16.88		14.87		12.45					
Percent Fantasy	WH	7.03		3.49		5.09		6.17	4.84		
	CS			2.77				5.03		9.10	3-5*
	WS	12.22		8.94		9.06					
Percent Explanation	WH	1.49		1.64		1.12		1.99	3.02		
	CS			1.70				1.72		2.76	
	WS	2.65		1.95	*	5.36					2½-3½**

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Limits	WH	2.03		1.02	*	4.02	**	0	1.93		
	CS			.64				.78		.92	
	WS	2.51		1.20	*	3.96					
Percent Clarification	WH	17.74		17.76	**	8.54		7.82	5.98		2½-4½*
	CS			13.08		**		7.36	*	4.82	3-5**
	WS	9.96	*	4.68		4.70					2½-3½*
Percent Description	WH	28.79		37.50		42.99		48.47	41.66		2½-3½* 2½-4½**
	CS			38.16				43.80		43.20	
	WS	37.26		45.22		41.13					
Percent Feeling	WH	1.44		2.57		2.38	**	1.27	2.51		
	CS			3.16				1.90		2.78	
	WS	3.22		3.42		2.83					
Percent Demonstration	WH	1.37		.48		.84		.45	.37		
	CS			.28				.36		.44	
	WS	.69		.81		.22					
Percent Commands	WH	1.52	*	.78		.06		0	.06		
	CS			.48				0		.14	
	WS	.26		.27		.15					
Percent Labels	WH	29.64		24.48		26.51		16.88	20.26		2½-4½*
	CS			25.12				19.62		21.59	
	WS	25.25	*	20.85		20.77					

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Specifies	WH	13.89		12.44		12.79		20.38	21.89		
	CS			14.59		*		21.08		20.52	3-5*
	WS	16.06		17.57		17.80					
Percent Utterances Responded To With Direct Information	WH	17.02		17.89	*	14.56	**	24.51	26.47		3½-4½**
	CS			21.25				20.70		20.84	
	WS	17.57		16.02	**	22.47					
Percent Utterances Responded To With Peripheral Information	WH	7.97	**	2.57		2.32		2.07	2.11		2½-3½** 2½-4½**
	CS			2.30				2.48		2.64	
	WS	6.15	**	1.69	**	3.27					2½-3½*
Percent Of Utterances Responded To With No Information	WH	14.06	*	18.77		14.61		20.47	14.06		
	CS			16.14				14.94		11.28	
	WS	* 16.77		10.95		12.98					
Percent Responded To	WH	39.11		39.32	*	31.59	**	47.20	42.74		2½-3½* 3½-4½**
	CS			39.86				38.18		34.92	
	WS	40.01	**	28.58	**	38.90					

Variable	Group	AGE 2½	3		3½		4	4½	5	Additional Significant Differences
Percent Ignored	WH	4.20	3.84	*	6.33	**	.76	.36		2½-3½* 3½-4½** 2½-4½*
	CS		3.68		*		1.16	*	1.32	
	WS	3.16	4.11		.86					2½-3½*
Percent Questions With Permanent Clarifying Information	WH	1.21	3.81		1.18		0	1.50		
	CS		2.51		*		.34		.52	
	WS	6.79	2.05		1.78					
Percent Questions With Transient Clarifying Information	WH	36.67	31.24	*	22.01		21.10	15.08		2½-4½*
	CS		26.26				21.08		14.94	3-5*
	WS	15.07	11.92		9.82					
Percent Statement With Permanent Clarifying Information	WH	.41	1.60		.12		0	.09		
	CS		.81				.02		.05	
	WS	3.09	.64		.42					

Variable	Group	AGE 2½		3		3½		4		4½		5	Additional Significant Differences	
Percent Statement With Transient Clarifying Information	WH	7.71		9.77	**	3.24		3.69		3.12			2½-3½*	2½-4½**
	CS			6.34		*		3.07		*		1.34	3-5**	
	WS	3.72		1.00		1.33								
Percent Question Directly Responded To	WH	37.8		38.73	*	31.62	**	51.00		52.64			3½-4½**	2½-4½**
	CS			40.24				44.42				42.84		
	WS	28.36		32.62		37.14								
Percent Questions Peripherally Responded To	WH	11.17	**	6.00		5.24	*	9.63	*	7.49			2½-3½**	
	CS			6.52				9.80				9.88		
	WS	9.92		5.15	**	9.45								
Percent Questions Responded To With No Information	WH	17.49	*	25.86	*	17.89		21.94		19.77				
	CS			23.72				19.39				18.26		
	WS	31.11		22.81		22.52								
Percent Statement Directly Responded To	WH	13.98		15.24		12.32	**	19.71		19.62			3½-4½*	
	CS			19.21				17.05				17.86		
	WS	18.07		14.45	*	21.08								

Variable	Group	AGE 2½		3		3½		4		4½	5	Additional Significant Differences
Percent Statement Peripheral Responded To	WH	7.81	**	1.46		.86		.39		.63		2½-3½** 2½-4½**
	CS			1.50				.70			.70	
	WS	5.53		.68		.92						2½-3½**
Percent Statement Responded To With No Information	WH	14.72		17.72		13.94	*	24.40	**	15.69		
	CS			16.31				16.80			11.00	
	WS	14.94		9.73		12.58						

① Asterisks between columns indicate the significance level of the difference between the means in those columns. (* = < .05; ** = < .01).

Appendix D
Table 42
Age Comparisons - Mothers

West Harlem (WH), Harlem Cross-Sectional (CS) and Washington Square (WS)
Mother Means and Mean Differences[Φ] for Age Comparisons on VINEX Variables

Variable	Group	AGE 2½		3		3½	4		4½	5	Additional Significant Differences	
Individual Utterance Frequency	WH	465.22		416.89	*	327.89	345.89	*	299.22		2½-3½**	2½-4½**
	CS			419.35		*	325.45			291.70	3-5**	
	WS	392.40	**	348.20		350.30						
Coded UH Frequency	WH	410.00		386.44	*	298.56	311.67		279.00			
	CS			379.35			294.35			260.55		
	WS	354.60	*	322.80		321.00						
Question Frequency	WH	147.11		180.00	*	132.00	110.33		102.56		2½-4½**	
	CS			137.65		*	94.55			75.85	3-5**	
	WS	133.30		112.10		114.20						
Statement Frequency	WH	262.89	*	205.56		166.56	201.67		176.44		2½-3½**	2½-4½**
	CS			241.65			199.95			184.75		
	WS	221.30		210.70		206.70						
Percent Coded	WH	88.78	*	92.29		91.37	89.37		93.64		2½-4½**	
	CS			90.42			89.57			89.17		
	WS	90.55		92.34		91.61						
Percent Questions	WH	36.74		46.32		43.84	34.71		37.59			
	CS			35.85			31.38			25.52	3-5*	
	WS	38.73		34.79		36.20						

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Statement	WH	63.16		53.61		56.06		65.34	62.28		
	CS			64.00				68.59		72.40	3-5*
	WS	61.17		65.11		63.67					
Percent Behavior Requests	WH	17.20	*	13.29		11.48		13.41	9.73		2½-3½** 2½-4½**
	CS			16.92				16.72		15.21	
	WS	12.63		13.25	*	9.19					
Percent Permanent Information	WH	36.18		37.51		35.16		31.58	34.69		
	CS			34.02				33.17		34.30	
	WS	43.19		34.73		37.89					
Percent Transient Information	WH	28.27	*	34.53		39.20	*	31.22	37.00		2½-3½*
	CS			34.05				29.28		31.23	
	WS	26.27	*	37.03		35.71					
Percent Behavioral Information	WH	31.07	**	23.31		20.40	**	30.58	* 23.56		2½-3½**
	CS			28.36				33.14		30.04	
	WS	20.50		23.10		21.43					
Percent Fantasy	WH	3.70		3.04		4.37		4.93	3.09		
	CS			2.04				3.08		3.36	
	WS	7.78		4.21		3.45					
Percent Explanation	WH	2.26		2.38		3.04		4.01	4.31		2½-4½**
	CS			3.10				3.62		4.50	
	WS	3.17	**	6.35		6.72					2½-3½**

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Limits	WH	4.20	*	2.17		3.51	**	0	1.91		2½-4½**
	CS			1.53				1.06		1.43	
	WS	4.06	*	1.66		3.29					
Percent Clarification	WH	14.54		17.72	**	9.67		7.27	6.36		2½-4½*
	CS			11.72		*		6.66		4.34	3-5**
	WS	10.96	*	5.13		4.53					2½-3½**
Percent Description	WH	35.48	*	43.67		50.16		55.39	50.32		2½-3½** 2½-4½**
	CS			45.14				50.82		51.14	
	WS	44.50		52.15		53.47					2½-3½*
Percent Feeling	WH	3.14		3.96	**	1.97	*	1.09	1.44		2½-3½* 2½-4½*
	CS			3.13		**		1.36		2.08	3-5*
	WS	2.87	**	1.04	*	2.25					
Percent Demonstration	WH	2.21	*	.24		.50		1.54	* .56		2½-3½** 2½-4½**
	CS			1.09				1.46		1.40	
	WS	1.37		1.37		.55					
Percent Commands	WH	1.47		.91		.32		.02	.18		2½-3½*
	CS			.73				.54		.30	
	WS	.52		.20		.19					
Percent Labels	WH	22.71		21.08		20.22		17.26	17.79		
	CS			20.61				18.01		18.14	
	WS	20.90		17.37		17.19					

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Specifies	WH	13.78	*	7.41		10.51		11.61	14.33		
	CS			10.60				14.26		14.86	3-5*
	WS	9.37	*	13.65	*	10.04					
Percent Utterances Responded To With Direct Information	WH	12.86	**	21.66	**	14.67	**	28.96	25.73		$3\frac{1}{2}$-$4\frac{1}{2}$** $2\frac{1}{2}$-$4\frac{1}{2}$**
	CS			20.88				24.07		22.36	
	WS	17.31		18.59		22.04					
Percent Utterances Responded To With Peripheral Information	WH	4.92	**	2.94		2.88		1.68	3.22		$2\frac{1}{2}$-$3\frac{1}{2}$* $2\frac{1}{2}$-$4\frac{1}{2}$*
	CS			2.00				2.62		2.94	
	WS	5.40	**	1.46		2.65					
Percent Of Utterances Responded To With No Information	WH	16.00		11.33		12.46		13.21	11.36		
	CS			9.90				12.64		11.42	
	WS	8.77		6.24		9.46					
Percent Responded To	WH	33.91		36.00		30.09	**	43.97	40.43		$3\frac{1}{2}$-$4\frac{1}{2}$*
	CS			32.88		*		39.44		36.92	
	WS	31.67		26.29	**	34.25					

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Ignored	WH	8.74		12.34		15.47	**	2.41	1.40		3½-4½** 2½-4½**
	CS			8.48		**		2.56		1.66	3-5**
	WS	9.33		6.21	*	1.42					2½-3½**
Percent Questions With Permanent Clarifying Information	WH	1.04		4.53	*	1.32		0	.48		
	CS			2.13				.52		.54	
	WS	9.02		2.29		2.53					
Percent Questions With Transient Clarifying Information	WH	24.29		22.66		17.91		15.92	11.78		2½-4½*
	CS			21.08				15.14		11.32	3-5**
	WS	13.85	**	9.89		6.25					2½-3½*
Percent Statement With Permanent Clarifying Information	WH	.40		.69		.18		0	.09		2½-4½*
	CS			.52				.26		.03	
	WS	1.35		.16		.22					

Variable	Group	AGE 2½		3		3½		4	4½	5	Additional Significant Differences
Percent Statement With Transient Clarifying Information	WH	6.54		7.09	**	1.18		2.73	2.50		2½-3½*
	CS			3.75				2.08		1.67	3-5*
	WS	1.44		.62		.95					
Percent Question Directly Responded To	WH	26.50	**	39.30	**	30.47	**	56.19	49.97		3½-4½** 2½-4½**
	CS			41.05		*		50.10		47.30	
	WS	31.34		38.50		44.38					2½-3'**
Percent Questions Peripherally Responded To	WH	6.03	*	4.20		4.72		4.73	6.39		
	CS			3.49				5.08		5.58	
	WS	9.28	**	4.38		4.48					2½-3½*
Percent Questions Responded To With No Information	WH	14.34		11.17		12.74		12.01	13.21		
	CS			11.51				14.86		14.70	
	WS	11.09		10.54		14.92					
Percent Statement Directly Responded To	WH	6.19		9.04		5.46	**	18.90	14.00		3½-4½** 2½-4½**
	CS			12.44				15.25		15.08	
	WS	8.12		9.08	*	12.07					2½-3½*

Variable	Group	AGE 2½		3	3	4		4½	5	Additional Significant Differences
Percent Statement Peripheral Responded To	WH	3.99	**	1.59	2.11	.29		1.74		2½-4½*
	CS			1.09		1.74			2.02	
	WS	3.39	**		.31	1.58				
Percent Statement Responded To With No Information	WH	18.01		12.26	14.19	15.08	*	11.69		
	CS			10.18		13.04			11.65	
	WS	8.26	*	4.62	7.63					

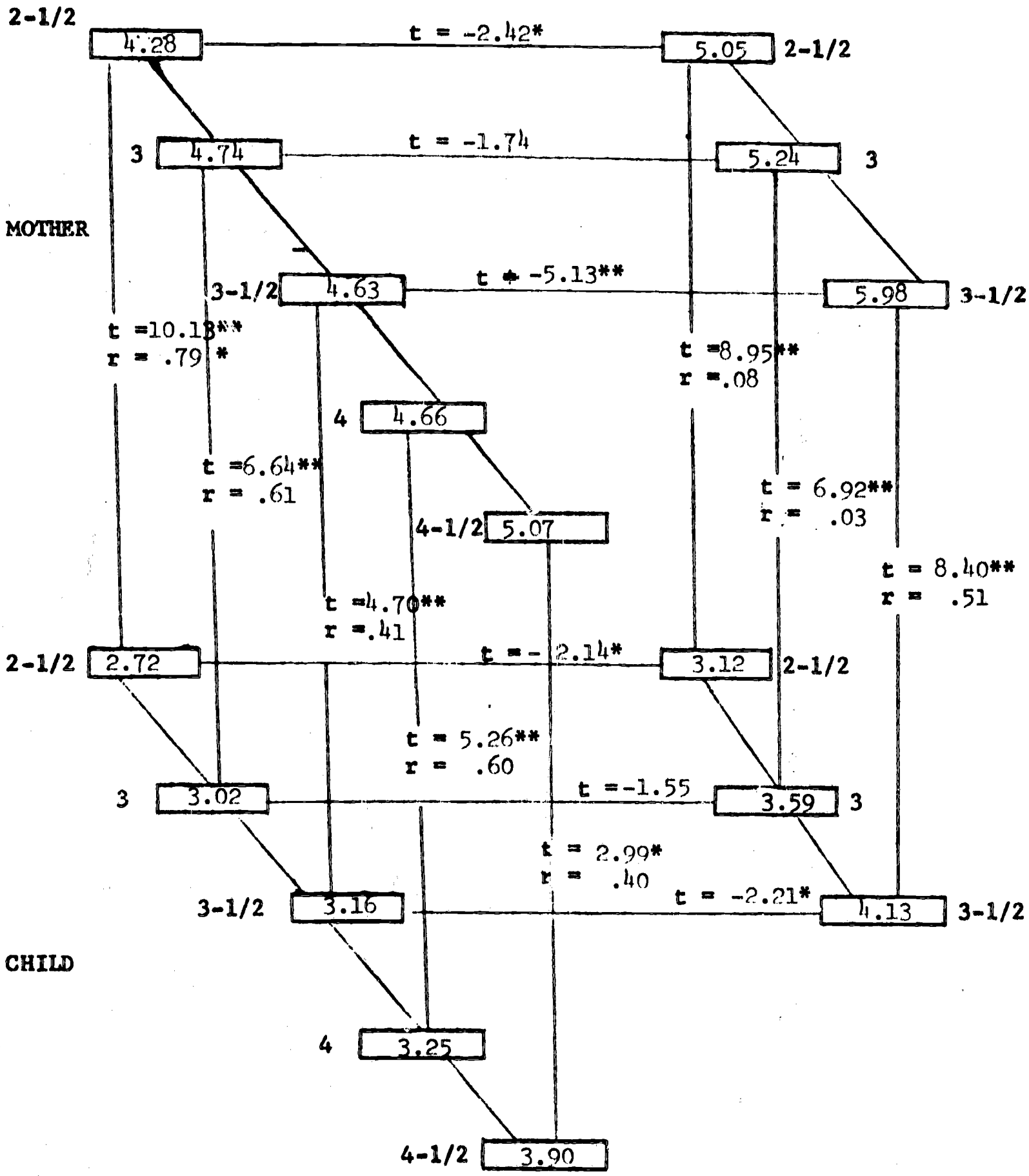
Asterisks between columns indicate the significance level of the difference between the means in those columns. (* = $p < .05$; ** = $p < .01$).

Appendix E

Sample and Age Differences - Syntactic Complexity

Appendix E
Table 1

Variable: Mean Complexity Scores by Age: Mothers and Children



Appendix E

Table 2

Mean Scores for all groups on five complexity measures

Mothers

Variable	Group	AGE 2½		3		3½	4		4½	5	Additional Significant Differences
Mean Words	WH	3.62		3.72		3.75	3.87		3.96		2½-4½**
	CS			3.56		*	4.02			4.10	3-5*
	WS	4.33		4.40	**	5.10					2½-3½**
Mean Complexity	WH	4.38	*	4.66		4.61	4.66	*	5.07		2½-3½* 2½-4½*
	CS			4.41			4.85			4.92	
	WS	5.03		5.20	**	5.98					2½-3½**
Number Different T's per Sentence	WH	.24	*	.26		.25	.28		.30		2½-3½* 3½-4½** 2½-4½**
	CS			.26		*	.29			.30	3-5*
	WS	.30		.32		.33					
Total Different T	WH	24.42	*	25.90		24.40	24.67	*	29.11		3½-4½** 2½-4½**
	CS			25.00			28.00			28.00	
	WS	30.40		32.50		34.10					2½-3½*
Number of T's per Sentence	WH	1.74		1.75		1.76	1.78		1.94		
	CS			1.62			1.78			1.80	
	WS	1.79		1.78		2.01					2½-3½*

Children

Variable	Group	AGE 2½		3		3½	4		4½	5	Additional Significant Differences		
Mean Words	WH	2.08		2.51		2.62	2.87		3.37		2½-3½**	3½-4½*	2½-4½**
	CS			2.52	*		3.03			3.43	3-5**		
	WS	2.67	**	3.08	**	3.56					2½-3½**		
Mean Complexity	WH	2.63		3.05		3.22	3.35	*	3.75		3½-4½*	2½-4½**	
	CS			2.99			3.41		*	3.99	3-5**		
	WS	3.13		3.56		4.13					3½-4½**		
Number Different T's per Sentence	WH	.15	*	.17		.19	.21	*	.24		2½-3½**	3½-4½*	2½-4½**
	CS			.18	**		.22		*	.29	3-5**		
	WS	.20	*	.22		.25					2½-3½**		
Total Different T	WH	15.08		17.30	*	19.10	20.44	*	24.44		2½-3½**	3½-4½**	2½-4½**
	CS			.18	**		.22			.25	3-5**		
	WS	19.73	*	21.91		24.30					2½-3½**		
Number of T's per Sentence	WH	1.14		1.12		1.15	1.13		1.36				
	CS			1.08			1.20			1.36	3-5**		
	WS	.96	*	1.20		1.47					2½-3½**		

Appendix F

Child Measures and Mother Interview

Appendix F

FIRST MOTHER INTERVIEW

Child's name
Birth date
Brother and sisters: names and ages
Adults living with family
Occupation of head of household

We are very glad you and _____ could come in, Mrs. _____. As we said before, we are interested in the way children learn about their world and how they change as they grow older. The best way to do this is to have them do something quite natural like playing with another person and recording what sorts of chings they say and do.

I'd like to ask you a few questions about the way his language has developed.

Mothers assessment of child's behavior in playroom.

1. Did it seem to you that _____ played the same way here that he does ordinarily or did he seem to be different in some way, maybe quieter or more excited or something?

Language Development

1. We can learn a good deal from listening to a child talk in a playroom like this, but we would also like to ask you more about the way _____ talks and plays at home. You know much more about him than we could ever learn in a short time like this.

2. When did _____ start saying his first words?

3. Do you remember what words he said?

4. Some children seem to say just one or two words at a time, but you always know what they mean. Other children babble away and you can swear they are saying something, but you just can't understand their babbling. It sounds just like a long sentence. Has _____ ever talked that way?

5. I expect that _____talks a lot more now than he did even a few months ago. Does he use any new words lately? Can you tell me some of them?

6. Do you notice any difference in the length of his sentences? What is the longest sentence you can remember him ever saying?

7. Some children like to use made-up words a lot, a kind of private language. Does _____ do that much now? Does he have any special names for things?

8. Are there any words or sounds that he seems to have difficulty with?

9. Children often have a lot of trouble mastering forms of grammar that adults use. Have you noticed whether _____ has particular trouble with some kinds of sentences or questions? Does he ever get the order of the words tangled up?

10. Many children like to make believe about themselves. How about _____? Does he ever pretend he's an animal, or somebody from television, or another member of the family?

Use of Grammatical Forms

11. Is he asking many questions these days?

12. Does he ask what the names of things are?

13. Does he ask where things are, or if something will happen?

14. Does he ask if he can do things?

15. Does he ever ask why he can't do things?

16. Does he ever ask other kinds of "why" questions?

Memory and anticipatory acts

17. Now let me ask about something a little different.
Does _____ ever show that he remembers things that happened yesterday or last week or a long time ago?
How far back does he seem to remember?
Does he ever show signs that he is expecting something to happen; the other children to come home from school or for dinner to be ready or something? Can you give me an example?
Does he ever seem to figure out what is happening even when he can't actually see it? When someone knocks on the door, when he hears the water running or something like that? Example?

Language interaction in family and with peers

18. When he talks, is it mostly to you or does he talk to (his sibs -- use names) the same way? Do they understand him as well as you do? Does he understand them as well as he does you?

19. Do they ever tell him about things, or teach him something? Example?

20. Is _____ at all interested in letters of the alphabet or numbers, or in figuring out what they mean? For instance, is he interested in the names on cereal boxes, or channel numbers on TV? Attending any play groups?

Play

21. Now just one more thing about his play. Does he have any toys at home now that are the same as the ones here?

22. What are some of his other toys?

23. Does he play a lot with them?

24. Does he play with one toy for a long time, or does he move from one to another quite fast?

Those are all the things I'd like to ask about, but perhaps you have some questions about the work we're doing here. Is there anything you'd like to know about our study?

Appendix F

Child's Name_____
Age_____
Date_____
Session_____
Exp_____

Imitation Task Selected Sentences

Frank & Osser[1]

Practice Sentence	Child's exact response
1. The bunny is eating a carrot.	_____
Test Sentences	
1. The boy is <u>not</u> on the chair.	_____
2. The boy is pulling the <u>girl's hair</u>.	_____
3. The boy dries <u>himself</u> with the towel.	_____
4. The <u>little</u> boy is flying the kite.	_____
5. The boy <u>who sits</u> is very fat.	_____
6. Mother does <u>some sweeping</u> with a broom.	_____
7. The girl sees <u>that the boy</u> sits.	_____

*Critical structure is underlined.

[1] The entire measure may be found in Osser, Wang, and Zaid (1969), p. 1065.

Appendix F

Berko's "WUGS" Test[1]

	Word Presented	Required Response
1.	wug	wugs
2.	rick	ricked
3.	mot	motting
4.	tass	tasses
5.	heaf	heafs or heaves
6.	glass	glasses
7.	gling	glinging
8.	loodge	loodges
9.	ring	rang
10.	wug	wug's
11.	zib	zibbing
12.	melt	melted
13.	wugs	wugs'

[1]. The complete measure may be found in Berko (1958).

Appendix F

Child's Name:_____ Date:_____

Session:_____

Observer:_____

Immediately after each session, rate the child's total performance; the mother's behavior and the interaction on the following dimensions. Please don't discuss your rating with other observers until after you have completed the ratings.

OBSERVER RATING SCALES

Child Characteristics

1. Body Build (check one)
   - chubby
   - average
   - slender

2. Motor coordination (check one)
   - poor
   - average
   - excellent
   - small
   - large

3. Level of language development (check one)
   - poor
   - average
   - excellent

   vocabulary size and quality articulation

   typical sentence structure

seldom more than 2 words in an utterances	uses several words but telegraphic	uses long grammatical sentences
1	2	3

On scales 4-13, place a check anywhere along the line.

4. Activity level

1	2	3
hyper-active	moderately active	moves about room very little

5. Child's interest level in play activities

1	2	3
wanders around room, plays little with materials. May or may not talk to mother	spends about half his time playing with materials	spends nearly all his time playing with materials

6. Interest span

1	2	3
flits from one material or activity	may move but returns to achieve and take up where he left off	tends to stay with one activity a relatively long time; or persistently returns to it

7. Amount of child's verbalization (may be relevant, irrelevant or even nonsense)

1	2	3
says almost nothing. during session may be involved with toy and not talking or may just	talks freely and spontaneously child not hesitant to talk, but may play for periods without talking	verbalizes or vocalizes most of time as he plays; chatters constantly sometimes even while mother is talking

8. Mother's interest in child's activities

1	2	3
shows very little interest, sits in mother's chair; may look out window or at self in mirror; seems bored much of time.		shows strong interest most of session-sits near him. watches his activities; comments frequently. She may not actively participate but is always interested.

9. Amount of mother's intervention

1	2	3
guides child's activity very little - even when want to or needs it	intervenes when child needs help or when requested but does not intrude	guides child's activity most of the time - either verbally or non-verbally

10. Amount of mother's own involvement with activities

1	2	3
Mother plays with toys - not at all	mother plays with toys occasionally - often the child asks her	mother actively plays with toys or verbally fantasies with child about them most of time

11. Mother's warmth

1	2	3
tone of voice cool or annoyed; lacks interest in what he is doing, or seems often dis-satisfied with it.		use warm tone of voice; responds to child's request and comments supportively shows interest in his activities; rewards child appropriately with smiles, praise, obviously enjoys being with child

12. Interaction

1	2	3
Mother and child talk very little while they are in room	Conversation is going on about half the time mother and child are together	Mother and child talk nearly all the time they are in the room

13. Emotional tone of session

1	2	3
Mother and child seem irritable with situation or each other	treat time as a mutually shared time both give ideas, make suggestions, respond to ideas of other	Both seem to enjoy session plus each others company; their relationship seems comfortable tone of voices generally happy.

Bibliography

Baldwin, A. L. Theories of Child Development. New York: Wiley, 1967.

Baldwin, A. L., & Frank, S. M. Syntactic complexity in mother-child interactions. Paper presented at the meeting of the Society for Research in Child Development, Santa Monica, March 1969.

Baldwin, C.P., & Baldwin, A. L. Children's judgments of kindness. Child Development, 1970, 41, 29-47.

Bereiter, C., & Englemann, S. Teaching disadvantaged children in pre-school. New York: Prentice-Hall, 1966.

Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-177.

Bernstein, B. Social structure, language, and learning. In J. P. De Cecco (Ed.), The psychology of language, thought, and instruction. New York: Holt, Rinehart, & Winston, 1967. Pp. 89-103.

Bijou, S. W. Child behavior and development: A behavioral analysis. International Journal of Psychology, 1968, 3, 221-238.

Billingsley, A. Black families in white America. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

Chomsky, N. Syntactic structures. The Hague: Mouton & Co., 1957.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: M. I. T. Press, 1965.

Culicover, P., Kimball, J., Lewis, C., Loveman, D., & Moyne, J. An automated recognition grammar for English. Technical Report, July 1969, International Business Machines Corp., FSC 69-5007.

Davis, W. A., & Havinghurst, R. J. Social class and color difference in child-rearing. American Sociological Review, 1946, 11, 698-710.

Deutsch, M. The role of social class in language development and cognition. American Journal of Orthopsychiatry, 1965, 35, 78-88.

Frank, S. , & Osser, H. A psycholinguistic model of syntactic complexity. Language and Speech (in press, 1970)

Harris, F. R., Wolf, M. M., & Baer, D. M. Effects of adult social reinforcement on child behavior. In W. W. Hartup & N. L. Smothergill (Eds.), The Young Child, Washington, D. C.: National Association for the Education of Young Children, 1967.

Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.

Hess, R. D., & Baer, R. M. (Eds.) Early Education. Chicago: Aldine Publishing Co., 1968.

Hess, R. D., Shipman, V. C., Brophy, J. E., & Baer, R. M. The cognitive environments of urban preschool children. O.E Final Report, Graduate School of Education, University of Chicago, 1968.

Hunt, J. McV. The psychological basis for using preschool enrichment as an antidote for cultural deprivation. Merrill-Palmer Quarterly of Behavior and Development, 1964, 10, 209-248.

Labov, W., & Cohen, P. Some suggestions for teaching standard English to speakers of non-standard urban dialects. Unpublished manuscript, 1968.

Levine, S. Stimulation in infancy. Scientific American, 1960, 202, 80-86.

Loban, W. Problems in oral English. Champaign, Ill.: National Council of Teachers, 1966.

McCaffrey, A. Communicative competence and the disadvantaged child: A study of the relationship between language models and communication behavior in disadvantaged preschoolers. Unpublished manuscript, May 1968.

Menyuk, P. Syntactic rules used by children from preschool through first grade. Child Development, 1964, 35, 533-546.

Milner, E. A study of the relationship between reading readiness in grade one school children and patterns of parent-child interaction. Child Development, 1951, 22, 95-112.

Moynihan, D. P. The Negro family: The case for national action. In L. Rainwater & W. L. Yancey. The Moynihan Report and the politics of controversy. Cambridge, Mass.: M. I. T. Press, 1967.

New York Times February 15, 1970.

Osser, H., Wang, M., & Zaid, F. The young child's ability to imitate and comprehend speech: A comparison of two subcultural groups. Child Development, 1969, 40, 1063-1075.

Reed, T. E. Caucasian Genes in American Negroes. Science, 1969, 165, 762-768.

Roberts, P. English Syntax. New York: Harcourt, Brace & World, 1964.

Rosenbaum, P. S. Specification and utilization of a transformational grammar. Scientific Report No. 2, October 1967, International Business Machines Corp., Contract AF 19 (628) - 5127, United States Air Force.

Salton, G. Automatic content analysis in information retrieval. Technical Report No. 68 - 5, January 1968, Department of Computer Science, Cornell University, Contract GN-495, National Science Foundation.

Scott, J. P. Comparative social psychology. In R. H. Waters, D. A. Rethlingshafer, & W. E. Caldwell (Eds), Principles of Comparative Psychology, New York: McGraw Hill, 1960.

Schank, R. C. A conceptual dependency representation for a computer-oriented semantics. Technical Report No. CS 130, March 1969, Computer Science Department, Stanford University, contract PHS MH 06645-07, National Institute of Mental Health.

Solomon, R. L. Punishment. American Psychologist, 1964, 19, 239-253.

Stewart, W. A. Urban Negro speech: Sociolinguistic factors affecting English teaching. In R. W. Shung (Ed.), Social dialects and language learning, Champaign, Ill.: National Council of Teachers of English, 1964.

Wooley, G. H. Syntactic analysis beyond the sentence. Paper presented at the fourth Annual meeting of the Association for Machine Translation and Computational Linguistics, Los Angles, July 1966.